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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
16 February 2017 (16.02.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/027397 A2

(51) International Patent Classification:
H04S 7/00 (2006.01) *G10L 25/78* (2013.01)
G10L 21/0272 (2013.01) *H04R 3/00* (2006.01)
H04R 1/10 (2006.01)

(21) International Application Number:
PCT/US2016/045834

(22) International Filing Date:
5 August 2016 (05.08.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

62/202,303	7 August 2015 (07.08.2015)	US
62/237,868	6 October 2015 (06.10.2015)	US
62/351,499	17 June 2016 (17.06.2016)	US
15/229,429	5 August 2016 (05.08.2016)	US

(71) Applicant: **CIRRUS LOGIC INTERNATIONAL SEMICONDUCTOR, LTD.** [GB/GB]; Westfield House, 26 Westfield Road, Edinburgh Lothian EH11 2QB (GB).

(72) Inventor; and
(71) Applicant *(for US only)*: **EBENEZER, Samuel Von Parma** [US/US]; 4630 S. Lakeshore Dr., Apt. 118, Tempe, Arizona 85282 (US).

(74) Agents: **PREWITT, Brian K.** et al.; Jackson Walker L.L.P., 100 Congress Avenue, Suite 1100, Austin, Texas 78701 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2017/027397 A2

(54) **Title**: EVENT DETECTION FOR PLAYBACK MANAGEMENT IN AN AUDIO DEVICE

(57) **Abstract**: In accordance with embodiments of the present disclosure, a method for processing audio information in an audio device may include reproducing audio information by generating an audio output signal for communication to at least one transducer of the audio device, receiving at least one input signal indicative of ambient sound external to the audio device, detecting from the at least one input signal a near-field sound in the ambient sound, and modifying a characteristic of the audio information reproduced to the at least one transducer in response to detection of the near-field sound.

EVENT DETECTION FOR PLAYBACK MANAGEMENT IN AN AUDIO DEVICE

Related Applications

The present disclosure claims priority to United States Non-Provisional Patent Application Serial No. 15/229,429, filed August 5, 2016, which claims priority to United States Provisional Patent Application Serial No. 62/202,303, filed August 7, 2015, United States Provisional Patent Application Serial No. 62/237,868, filed October 6, 2015, and United States Provisional Patent Application Serial No. 62/351,499, filed June 17, 2016, each of which is incorporated by reference herein in its entirety.

Field

The field of representative embodiments of this disclosure relates to methods, apparatuses, or implementations concerning or relating to playback management in an audio device. Applications include detection of certain ambient events, but are not limited to, those concerning the detection of near-field sound, proximity sound and tonal alarm detection using spatial processing based on signals received from multiple microphones.

Background

Personal audio devices have become prevalent and they are used in diverse ambient environments. The headphones used in these audio devices have become advanced such that the occlusion caused by either passive or active methods prevents a user from keeping track of an ambient sound field external to the audio device. Even though the increased isolation and uninterrupted listening is preferable in most cases, sometimes for safety or enhanced user experience, it is imperative that some specific ambient events are heard by the user and an appropriate action is taken in response to that event. For example, if the user is listening to music through his headset and interrupted by someone attempting to start a conversation with him or her, it may be difficult to maintain the conversation unless the user pauses the playback signal or reduces the volume of the playback signal. For example, U.S. Pat. No. 7,903,825

proposes an audio device in which the playback signal is modified depending on the ambient acoustic field. As another example, U.S. Pat. No. 8,804,974 teaches ambient event detection in a personal audio device which can then be used to implement an event-based modification of the playback content. The above-mentioned references also teach the use of microphones to detect various acoustic events. As a further example, U.S. App. Ser. No. 14/324,286, filed on July 7, 2014, teaches using a speech detector as an event detector to adjust the playback signal during a conversation. As an additional example, U.S. Pat. No. 8,565,446 teaches the use of a direction of arrival (DOA) estimate and an interference to desired (near-field) speech signal ratio estimate from a set of plural microphones to detect desired speech in the presence of non-stationary background noise to control a speech enhancement algorithm in a noise reduction echo cancellation (NREC) system. Similarly, U.S. App. Ser. No. 13/199,593 teaches that a maximum of the normalized cross-correlation statistic that is derived through a cross-correlation analysis of plural microphones may be an effective discriminator to detect near-field speech. A spectral flatness measure-based music detector for a NREC system is proposed in U.S. Pat. No. 8,126,706 to differentiate the presence of background noise from the background music. U.S. Pat. Nos. 7,903,825; 8,804,974; U.S. App. Ser. No. 14/324,286; U.S. Pat. No. 8,565,446; U.S. App. Ser. No. 13/199,593; and U.S. Pat. No. 8,126,706 are incorporated by reference herein.

Summary

In accordance with the teachings of the present disclosure, one or more disadvantages and problems associated with existing approaches to event detection for playback management in a personal audio device may be reduced or eliminated.

In accordance with embodiments of the present disclosure, a method for processing audio information in an audio device may include reproducing audio information by generating an audio output signal for communication to at least one transducer of the audio device, receiving at least one input signal indicative of ambient sound external to the audio device, detecting from the at least one input signal a near-field sound in the ambient sound, and modifying a characteristic of the audio information reproduced to the at least one transducer in response to detection of the near-field sound.

In accordance with these and other embodiments of the present disclosure, an integrated circuit for implementing at least a portion of an audio device may include an audio output configured to reproduce audio information by generating an audio output signal for communication to at least one transducer of the audio device, a microphone input configured to receive an input signal indicative of ambient sound external to the audio device, and a processor configured to detect from the input signal a near-field sound in the ambient sound and modify a characteristic of the audio information in response to detection of the near-field sound.

In accordance with these and other embodiments of the present disclosure, a method for processing audio information in an audio device may include reproducing audio information by generating an audio output signal for communication to at least one transducer of the audio device, receiving at least one input signal indicative of ambient sound external to the audio device, detecting from the at least one input signal an audio event, and modifying a characteristic of the audio information reproduced to the at least one transducer in response to detection of the audio event being persistent for at least a predetermined time.

In accordance with these and other embodiments of the present disclosure, an integrated circuit for implementing at least a portion of an audio device, may include an audio output configured to reproduce audio information by generating an audio output signal for communication to at least one transducer of the audio device, a microphone input configured to receive an input signal indicative of ambient sound external to the audio device, and a processor configured to detect from the input signal an audio event and modify a characteristic of the audio information reproduced to the at least one transducer in response to detection of the audio event being persistent for at least a predetermined time.

Technical advantages of the present disclosure may be readily apparent to one of ordinary skill in the art from the figures, description and claims included herein. The objects and advantages of the embodiments will be realized and achieved at least by the elements, features, and combinations particularly pointed out in the claims.

It is to be understood that both the foregoing general description and the following detailed description are examples and explanatory and are not restrictive of the claims set forth in this disclosure.

Brief Description of the Drawings

A more complete understanding of the example, present embodiments and certain advantages thereof may be acquired by referring to the following description taken in conjunction with the accompanying drawings, in which like reference numbers indicate like features, and wherein:

Figure 1 illustrates an example of a use case scenario wherein such detectors may be used in conjunction with a playback management system to enhance a user experience, in accordance with embodiments of the present disclosure;

Figure 2 illustrates an example playback management system that modifies a playback signal based on a decision from an event detector, in accordance with embodiments of the present disclosure;

Figure 3 illustrates an example event detector, in accordance with embodiments of the present disclosure;

Figure 4 illustrates functional blocks of a system for deriving near-field spatial statistics that may be used to detect audio events, in accordance with embodiments of the present disclosure;

Figure 5 illustrates example fusion logic for detecting near-field sound, in accordance with embodiments of the present disclosure;

Figure 6 illustrates example fusion logic for detecting proximity sound, in accordance with embodiments of the present disclosure;

Figure 7 illustrates an embodiment of a proximity speech detector , in accordance with embodiments of the present disclosure;

Figure 8 illustrates example fusion logic for detecting a tonal alarm event, in accordance with embodiments of the present disclosure;

Figure 9 illustrates an example timing diagram illustrating hold-off and hang-over logic that may be applied on an instantaneous audio event detection signal to generate a validated audio event signal, in accordance with embodiments of the present disclosure; and

Figure 10 illustrates different audio event detectors having hold-off and hang-over logic, in accordance with embodiments of the present disclosure.

Detailed Description

In accordance with embodiments of this disclosure, systems and methods may use at least three different audio event detectors that may be used in an automatic playback management framework. Such audio event detectors for an audio device may include a near-field detector that may detect when sounds in the near-field of the audio device is detected, such as a user of the audio device (e.g., a user that is wearing or otherwise using the audio device) speaks, a proximity detector that may detect when sounds in proximity to the audio device is detected, such as when another person in proximity to the user of the audio device speaks, and a tonal alarm detector that detects acoustic alarms that may have been originated in the vicinity of the audio device are proposed. Figure 1 illustrates an example of a use case scenario wherein such detectors may be used in conjunction with a playback management system to enhance a user experience, in accordance with embodiments of the present disclosure.

Figure 2 illustrates an example playback management system that modifies a playback signal based on a decision from an event detector 2, in accordance with embodiments of the present disclosure. Signal processing functionality in a processor 50 may comprise an acoustic echo canceller 1 that may cancel an acoustic echo that is received at microphones 52 due to an echo coupling between an output audio transducer 51 (e.g., loudspeaker) and microphones 52. The echo reduced signal may be communicated to event detector 2 which may detect one or more various ambient events, including without limitation a near-field event (e.g., including but not limited to speech from a user of an audio device) detected by near-field detector 3, a proximity event (e.g., including but not limited to speech or other ambient sound other than near-field sound) detected by proximity detector 4, and/or a tonal alarm event detected by alarm detector 5. If an audio event is detected, an event-based playback control 6 may modify a characteristic of audio information (shown as "playback content" in Figure 2) reproduced to

output audio transducer 51. Audio information may include any information that may be reproduced at output audio transducer 51, including without limitation, downlink speech associated with a telephonic conversation received via a communication network (e.g., a cellular network) and/or internal audio from an internal audio source (e.g., music file, video file, etc.).

Figure 3 illustrates an example event detector, in accordance with embodiments of the present disclosure. As shown in Figure 3, the example event detector may comprise a voice activity detector 10, a music detector 9, a direction of arrival estimator 7, a near-field spatial information extractor 8, a background noise level estimator 11, and decision fusion logic 12 that uses information from voice activity detector 10, music detector 9, direction of arrival estimator 7, near-field spatial information extractor 8, and background noise level estimator 11 to detect audio events, including without limitation, near-field sound, proximity sound other than near-field sound, and a tonal alarm.

Near-field detector 3 may detect near-field sounds including speech. When such near-field sound is detected, it may be desirable to modify audio information reproduced to output audio transducer 51, as detection of near-field sound may indicate that a user is participating in a conversation. Such near-field detection may need to be able to detect near-field sound in acoustically noisy conditions and be resilient to false detection of near-field sounds in very diverse background noise conditions (e.g., background noise in a restaurant, acoustical noise when driving a car, etc.). As described in greater detail below, near-field detection may require spatial sound processing using a plurality of microphones 51. In some embodiments, such near-field sound detection may be implemented in a manner identical or similar to that described in U.S. Pat. No. 8,565,446 and/or U.S. App. Ser. No. 13/199,593.

Proximity detector 4 may detect ambient sounds (e.g., speech from a person in proximity to a user, background music, etc.) other than near-field sounds. As described in greater detail below, because it may be difficult to differentiate proximity sounds from non-stationary background noise and background music, proximity detector may utilize a music detector and noise level estimation to disable proximity detection of proximity detector 4 in order to avoid

poor user experience due to false detection of proximity sounds. In some embodiments, such proximity sound detection may be implemented in a manner identical or similar to that described in U.S. Pat. No. 8,126,706, U.S. Pat. No. 8,565,446, and/or U.S. App. Ser. No. 13,199,593.

Tonal alarm detector 5 may detect tonal alarms (e.g., sirens) proximate to an audio device. To provide maximum user experience, it may be desirable that tonal alarm detector 5 ignores certain alarms (e.g., feeble or low-volume alarms). As described in greater detail below, tonal alarm detection may require spatial sound processing using a plurality of microphones 51. In some embodiments, such proximity sound detection may be implemented in a manner identical or similar to that described in U.S. Pat. No. 8,126,706 and/or U.S. App. Ser. No. 13,199,593.

Figure 4 illustrates functional blocks of a system for deriving near-field spatial statistics that may be used to detect audio events, in accordance with embodiments of the present disclosure. The level analysis 41 may be performed on microphones 52 by estimating the inter-microphone level difference (*imd*) between the near and far microphone (e.g., as described in U.S. App. Ser. No. 13/199,593). Cross-correlation analysis 13 may be performed on signals received by microphones 52 to obtain the direction of arrival information *DOA* of ambient sound that impinges on microphones 52 (e.g., as described in U.S. Pat. No. 8,565,446). In cross-correlation analysis 13, a maximum normalized correlation value *normMaxCorr* may also be obtained (e.g., as described in U.S. App. Ser. No. 13/199,593). Voice activity detector 10 may detect presence of speech and generate a signal *speechDet* indicative of present or absence of speech in the ambient sound (e.g., as described in the probabilistic based speech presence/absence based approach of U.S. Pat. No. 7,492,889). Beamformers 15 may, based on signals from microphones 52, generate a near-field signal estimate and an interference signal estimate which may be used by a noise analysis 14 to determine a level of noise *noiseLevel* in the ambient sound and an interference to near-field signal ratio *idr*. U.S. Pat. No. 8,565,446 describes an example approach for estimating interference to near-field signal ratio *idr* using a pair of beamformers 15. A voice activity detector 36 may use the interference estimate to detect (*proxSpeechDet*) any speech signal that does not originate from the desired signal direction.

Noise analysis 14 may be performed based on the direction of arrival estimate *DOA* by updating interference signal energy whenever the direction of arrival estimate *DOA* of the ambient sound is outside the acceptance angle of the near-field sound. The direction of arrival of the near-field sounds may be known *a priori* for a given microphone array configuration in the industrial design of a personal audio device.

The various statistics generated by the system of Figure 4 may then be used to detect the presence of near-field sound. Figure 5 illustrates example fusion logic for detecting near-field sound, in accordance with embodiments of the present disclosure. As shown in Figure 5, near-field speech may be detected when all the following criteria are satisfied:

- Direction of arrival estimate DOA of ambient sound is within an acceptance angle of near-field sound (block 16);
- Maximum normalized cross-correlation statistic *normMaxCorr* is greater than a threshold *normMaxCorrThres1* (block 17);
- Interference to near-field desired signal ratio *idr* is smaller than a threshold *idrThres1* (block 18);
- Voice activity is detected as indicated by signal *speechDet* (block 19); and
- Inter-microphone level difference statistic *imd* is greater than a threshold *imdTh* (block 42).

In some embodiments, thresholds *idrThres* and *imdTh* may be dynamically adjusted based on a background noise level estimate.

Proximity detection of proximity detector 4 may be different than near-field sound detection of near-field detector 3 because the signal characteristics of proximity speech may be very similar to ambient signals such as music and noise. Accordingly, proximity detector 4 must avoid false detection of proximity speech in order to achieve acceptable user experience.

Accordingly, a music detector 9 may be used to disable proximity detection whenever there is music in the background. Similarly, proximity detector 4 may be disabled whenever background noise level is above certain threshold. The threshold value for background noise may be determined *a priori* such that a likelihood of false detection below the threshold level is very low. Figure 6 illustrates example fusion logic for detecting proximity sound (e.g. speech), in accordance with embodiments of the present disclosure. Moreover, there may exist many environment noise sources that generate acoustic stimuli that are transient in nature. These noise types can be falsely detected as speech signal by the speech detector. To reduce the likelihood of false detection, a spectral flatness measure (SFM) statistic from the music detector 9 may be used to distinguish speech from transient noises. For example, the SFM may be tracked over a period of time and the difference between the maximum and the minimum SFM value over the same duration, defined as *sfmSwing* may be calculated. The value of *sfmSwing* may generally be small for transient noise signals as the spectral content of these signals are wideband in nature and they tend to be stationary for a short interval of time (300-500 ms). The value of *sfmSwing* may higher for speech signals because the spectral content of speech signal may vary faster than transient signals. As shown in Figure 6, proximity sound (e.g., speech) may be detected when all the following criteria are satisfied:

- Music is not detected in the background (block 20);
- Direction of arrival of estimate *DOA* is within an acceptance angle of proximity sound (block 21);
- Maximum normalized cross-correlation statistic *normMaxCorr* is greater than a threshold, *normMaxCorrThres2* (block 22);
- The background noise level *noiseLevel* is below a threshold *noiseLevelTh* (block 23); and
- Proximity voice activity is detected, as indicated by signal prox*SpeechDet* (block 19);

- SFM variation statistic *sfmSwing* is greater than a threshold *sfmSwingTh* (block 37);

- Interference to near-field desired signal ratio *idr* is greater than a threshold *idrThres2* (block 40); and

- Inter-microphone level difference statistic *imd* is close to 0 dB (block 43).

In some embodiments, the music detector taught in U.S. Pat. No. 8,126,706 may be used to implement music detector 9 to detect the presence of background music. Another embodiment of the proximity speech detector is shown in Figure 7, in accordance with embodiments of the present disclosure. According to this embodiment, proximity speech may be detected if the following conditions are met:

- Interference to near-field desired signal ratio *idr* is greater than a threshold *idrThres2* (block 39);

- Proximity voice activity is detected (block 27);

- Maximum normalized cross-correlation statistic *normMaxCorr* is greater than a threshold, *normMaxCorrThres3* (block 28);

- Direction of arrival of estimate *DOA* is within an acceptance angle of proximity sound (block 29);

- Music is not detected in the background (block 30);

- Low or medium level background or no background noise is present (block 31). This condition is verified by comparing the estimated background noise level with a threshold, *noiseLevelThLo*. If low noise level is detected, then the following two conditions are further tested to confirm the presence of proximity speech;

- SFM variation statistic *sfmSwing* is greater than a threshold *sfmSwingTh* (block 38);

- Inter-microphone level difference statistic *imd* is closer to 0 dB (block 44).

If the above-mentioned background noise level condition is not satisfied at block 31, then the following conditions may be indicative of proximity speech, in order to improve the detection rate of proximity speech without increasing occurrence of a false alarm (e.g., due to background noise conditions):

- Stationary background noise is present (block 32). The stationary background noise may be detected by calculating the ratio of peak-to-root mean square value of the SFM generated by music detector (block 9) over a period of time. Specifically, if the above-mentioned ratio is higher, then non-stationary noise may be present as the spectral flatness measure of a non-stationary noise tends to change faster than stationary noises;

- High noise level is present (block 32). The high noise-condition may be detected if the estimated background noise is greater than a threshold, *noiseLevelLo* and smaller than a threshold, *noiseLevelHi*.

If the above stationary noise and the direction of arrival conditions are not satisfied at block 32, then the presence of both of the following set of conditions may indicate the presence of proximity speech:

- Close-talking proximity talker is present (block 33). A close-talking proximity talker may be detected when the maximum normalized cross-correlation statistic *normMaxCorr* is greater than a threshold, *normMaxCorrThres4* (the threshold *normMaxCorrThres4* may be greater than *normMaxCorrThres3* to indicate the presence of close talker);

- Low- or medium- or high-level background or no background noise is present (block 34). This condition may be detected if the estimated background noise level is less than a threshold *noiseLevelThHi*.

If the above-mentioned direction of arrival condition is not satisfied at block 29, then the presence of following conditions may be indicative of proximity speech:

- The absence of music (block 35);

- Close-talking proximity talker is present (block 33). A close-talking proximity talker may be detected when the maximum normalized cross-correlation statistic *normMaxCorr* is greater than a threshold, *normMaxCorrThres4* (the threshold *normMaxCorrThres4* may be greater than *normMaxCorrThres3* to indicate the presence of close talker);

- Low- or medium- or high-level background or no background noise is present (block 34). This condition may be detected if the estimated background noise level is less than a threshold *noiseLevelThHi*.

Tonal alarm detector 5 may be configured to detect alarm signals that are tonal in nature in which a sonic bandwidth of such alarm signals are also narrow (e.g., siren, buzzer). In some embodiments, the tonality of an ambient sound may be measured by splitting the time domain signal into multiple sub-bands through time to frequency domain transformation and the spectral flatness measure, depicted in Figure 6 as signal *sfm[]* generated by music detector 9, may be computed in each sub-band. Spectral flatness measures *sfm[]* from all sub-bands may be evaluated, and a tonal alarm event may be detected if the spectrum is flat in most sub-bands but not in all sub-bands. Moreover, in a playback management system, it may not be necessary to detect far–field alarm signals. Accordingly, near-field spatial statistics 8 of Figure 3 may be used to differentiate the far-field alarm signals from near-field signals. Figure 8 illustrates example fusion logic for detecting a tonal alarm event (e.g. siren, buzzer), in accordance with

embodiments of the present disclosure. As shown in Figure 8, a tonal alarm event may be detected when all the following criteria are satisfied:

- Direction of arrival estimate *DOA* is within an acceptance angle of the alarm signal (block 24);
- Maximum normalized cross-correlation statistic *normMaxCorr* is greater than a threshold, *normMaxCorrThres5* (block 25); and
- Spectral flatness measure *sfm[]* indicated that the noise spectrum is flat in most sub-bands but not all (block 26).

In practice, the instantaneous audio event detections of near-field detector 3, proximity detector 4, and tonal alarm detector 5 as shown in Figure 5, 6, 7, and 8 may indicate false audio events. Accordingly, it may be desirable to validate an instantaneous audio event detection signal before communicating an event detection signal to playback control block 6. Figure 9 illustrates an example timing diagram illustrating hold-off and hang-over logic that may be applied on an instantaneous audio event detection signal to generate a validated audio event signal, in accordance with embodiments of the present disclosure. As shown in Figure 9, hold-off logic may generate a validated audio event signal in response to instantaneous detection of an audio event (e.g., near-field sound, proximity sound, tonal alarm event) being persistent for at least a predetermined time, while hang-over logic may continue to assert the validated audio event signal until the instantaneous detection of an audio event has ceased for a second predetermined time.

The following pseudo-code may demonstrate application of the hold-off and hang-over logic to reduce false detection of audio events, in accordance with embodiments of the present disclosure.

```
/* If the instant. detect is true, increment the hold off counter and reset the hang over
 * counter */
```

```
if(instDet == TRUE)
{
 holdOffCntr = holdOffCntr + 1;
 hangOverCntr = 0;
}
/* If the instant. detect is false, increment the hang over counter and reset the hold off
 * counter */
else
{
 hangOverCntr = hangOverCntr + 1;
 holdOffCntr = 0;
}
/******************
 * Hold-off Logic *
 ******************/
/* Valid detect will transition to true state if the instant. detect is continuously true for
 * certian time and the previous valid detect is false */
if(holdOffCntr > holdOffThres && validDet == FALSE)
{
```

```
 validDet = TRUE;
 holdOffCntr = 0;
 hangOverCntr = 0;
}
/*******************
 * Hang-Over Logic *
 *******************/
/* Valid NF detect will transition to false state if the instant. NF detect is continuously
 * false for certain time and the previous valid NF detect is true */
if(hangOverCntr > hangOverThres && validDet == TRUE)
{
 validDet    = FALSE;
 holdOffCntr  = 0;
 hangOverCntr = 0;
}
```

A validated event may be further validated before generating the playback mode switching control. For example, the following pseudo-code may demonstrate application of the hold-off and hang-over logic for gracefully switching between a conversational mode (e.g., in which audio information reproduced to output audio transducer 51 may be modified in response to an audio event) and a normal playback mode (e.g., in which the audio information reproduced to output audio transducer 51 is unmodified).

```
/***********************************
 * Conversational Mode Enter Logic *
 ***********************************/
/* Increment the time to enter conversational mode counter if the event detect is true
 * and the mode is not in the conversational mode.  If the counter exceeds the
 * threshold, switch to conversational mode and reset the counters.  Note that the
 * the event detect need not be true contiguously. */
if(convModeEn == FALSE && validDet == TRUE)
{
 timeToEnterConvModeCntr = timeToEnterConvModeCntr + 1;
 if(timeToEnterConvModeCntr > timeToEnterConvModeThres)
 {
  convModeEn = TRUE;
  timeToEnterConvModeCntr = 0;
  timeToExitConvModeCntr = 0;
 }
}
/***********************************
 * Conversational Mode Exit Logic *
```

```
**********/
/* Increment the time to exit conversational mode counter if the event detect is false
 * and the mode is in the conversational mode.  If the counter exceeds the
 * threshold, switch to normal mode and reset the counters.  Note that the
 * the event detect must be false contiguously. */
if(convModeEn == TRUE && validDet == FALSE)
{
 timeToExitConvModeCntr++;
 if(timeToExitConvModeCntr > timeToExitConvModeThres)
 {
  convModeEn = FALSE;
  timeToEnterConvModeCntr = 0;
  timeToExitConvModeCntr = 0;
 }
}
else
{
 timeToExitConvModeCntr = 0;
}
```

Figure 10 illustrates different audio event detectors having hold-off and hang-over logic, in accordance with embodiments of the present disclosure. The hold-off periods and/or hang-over periods for each detector may be set differently. In addition, in some embodiments, the playback management may be controlled differently based on the type of detected event. In these and other embodiments, as shown in Figure 9, a playback gain (and hence the audio information reproduced at output audio transducer 51) may be attenuated whenever one or more of the audio events is detected. In these and other embodiments, in order to provide smooth gain transition, a playback gain may be smoothed using a first order exponential averaging filter represented by the following pseudo-code:

```
if(convModeEn == TRUE)
{
 playBackGain = (1-alpha)*convModeGain + alpha*playBackGain
}
else
{
 playBackGain = (1-beta)*normalModeGain + beta*playBackGain
}
```

The smoothing parameters alpha and beta may be set at different values to adjust a gain ramping rate.

It should be understood—especially by those having ordinary skill in the art with the benefit of this disclosure—that that the various operations described herein, particularly in connection with the figures, may be implemented by other circuitry or other hardware components. The order in which each operation of a given method is performed may be changed, and various elements of the systems illustrated herein may be added, reordered, combined,

omitted, modified, etc. It is intended that this disclosure embrace all such modifications and changes and, accordingly, the above description should be regarded in an illustrative rather than a restrictive sense.

Similarly, although this disclosure makes reference to specific embodiments, certain modifications and changes can be made to those embodiments without departing from the scope and coverage of this disclosure. Moreover, any benefits, advantages, or solutions to problems that are described herein with regard to specific embodiments are not intended to be construed as a critical, required, or essential feature or element.

Further embodiments likewise, with the benefit of this disclosure, will be apparent to those having ordinary skill in the art, and such embodiments should be deemed as being encompassed herein.

WHAT IS CLAIMED IS:

1. A method for processing audio information in an audio device comprising:

reproducing audio information by generating an audio output signal for communication to at least one transducer of the audio device;

receiving at least one input signal indicative of ambient sound external to the audio device;

detecting from the at least one input signal a near-field sound in the ambient sound; and

modifying a characteristic of the audio information reproduced to the at least one transducer in response to detection of the near-field sound.

2. The method of Claim 1, further comprising determining from the at least one input signal a direction of the ambient sound and modifying the characteristic of the audio information reproduced to the at least one transducer in response to the direction of the ambient sound indicating that the ambient sound is sound from a user of the audio device.

3. The method of Claim 1, further comprising determining from the at least one input signal a direction of the ambient sound and modifying the characteristic of the audio information reproduced to the at least one transducer in response to the direction of the ambient sound indicating that the ambient sound is speech from a user of the audio device.

4. The method of Claim 1, wherein modifying the characteristic of the audio information comprises attenuating the audio information.

5. The method of Claim 1, further comprising modifying the characteristic of the audio information in response to a detection of the near-field sound being persistent for at least a predetermined time.

6. The method of Claim 5, further comprising:

detecting from the at least one input signal absence of the near-field sound in the ambient sound; and

ceasing to modify the characteristic of the audio information reproduced to the at least one transducer in response to the absence of the near-field sound for at least a second predetermined time.

7. The method of Claim 1, further comprising:

in addition to detecting the near-field sound, detecting from the at least one input signal ambient sound other than the near-field sound in the ambient sound; and

modifying the characteristic of the audio information reproduced to the at least one transducer in response to detection of the ambient sound.

8. The method of Claim 7, further comprising determining from the at least one input signal a direction of the ambient sound and modifying the characteristic of the audio information reproduced to the at least one transducer in response to the direction of the ambient sound indicating that the ambient sound is sound other than the near-field sound.

9. The method of Claim 7, further comprising:

detecting from the at least one input signal whether the ambient sound comprises background noise; and

modifying the characteristic of the audio information reproduced to the at least one transducer in response to detection of the background noise in the ambient sound.

10. The method of Claim 7, further comprising:

detecting from the at least one input signal whether the ambient sound comprises a tonal alarm; and

modifying the characteristic of the audio information reproduced to the at least one transducer in response to detection of the tonal alarm in the ambient sound.

11. The method of Claim 10, wherein detecting the tonal alarm in the ambient sound comprises:

detecting from the at least one input signal a direction of the ambient sound;

detecting from the at least one input signal a spectral flatness measure of the ambient sound;

detecting from the at least one input signal near-field spatial statistics of the ambient sound; and

detecting the tonal alarm based on the direction of the ambient sound, the presence or absence of the background noise, and the near-field spatial statistics.

12. The method of Claim 11, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise a correlation between the first microphone signal and the second microphone signal.

13. The method of Claim 11, wherein detecting the direction of the ambient sound comprises determining whether the direction of the ambient sound is within an acceptance angle of near-field sound.

14. The method of Claim 11, wherein detecting the near-field spatial statistics of the ambient sound comprises detecting whether a normalized cross-correlation statistic is greater than a threshold.

15. The method of Claim 11, wherein detecting the spectral flatness measure of the ambient sound comprises detecting whether the noise spectrum is flat in most sub-bands of the ambient sound, but not all sub-bands of the ambient sound.

16. The method of Claim 1, wherein detecting the near-field sound in the ambient sound comprises:

detecting from the at least one input signal a direction of the ambient sound;

detecting from the at least one input signal a presence of speech in the ambient sound;

detecting from the at least one input signal near-field spatial statistics of the ambient sound; and

detecting the near-field sound based on the direction, presence or absence of speech, and the near-spatial statistics of the ambient sound.

17. The method of Claim 16, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise a correlation between the first microphone signal and the second microphone signal.

18. The method of Claim 16, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise an interference-to-signal ratio associated with the near-field sound.

19. The method of Claim 16, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise an inter-microphone level difference between the first microphone signal and the second microphone signal.

20. The method of Claim 16, wherein detecting the direction of the ambient sound comprises determining whether the direction of the ambient sound is within an acceptance angle of near-field sound.

21. The method of Claim 16, wherein detecting the near-field spatial statistics of the ambient sound comprises:

detecting whether a normalized cross-correlation statistic is greater than a first threshold;

detecting whether an interference to near-field desired signal ratio is lesser than a second threshold; and

detecting whether an inter-microphone level difference is greater than a third threshold.

22. The method of Claim 21, wherein the second threshold is adjusted based on an estimate of background noise in the ambient sound.

23. The method of Claim 21, wherein the third threshold is adjusted based on an estimate of background noise in the ambient sound.

24. The method of Claim 1, further comprising:

detecting from the at least one input signal a direction of the ambient sound;

detecting from the at least one input signal a presence of background noise in the ambient sound;

detecting from the at least one input signal a presence of proximity speech in the ambient sound;

detecting from the at least one input signal a volume of the ambient sound;

detecting from the at least one input signal near-field spatial statistics of the ambient sound;

detecting based on the direction, presence or absence of background noise, presence or absence of the speech, the volume, and the near-spatial statistics of the ambient sound, a presence of an audio event comprising a proximity sound event; and

modifying the characteristic of the audio information reproduced to the at least one transducer in response to presence of the audio event.

25. The method of Claim 24, further comprising:

detecting variation in spectral content of the ambient sound; and

detecting based on the direction, presence or absence of background noise, presence or absence of the speech, the volume, the near-spatial statistics of the ambient sound, and the spectral content of the ambient sound, a presence of an audio event comprising a proximity sound event.

26. The method of Claim 25, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise a correlation between the first microphone signal and the second microphone signal.

27. The method of Claim 25, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise an interference-to-signal ratio associated with the near-field sound.

28. The method of Claim 25, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise an inter-microphone level difference between the first microphone signal and the second microphone signal.

29. The method of Claim 25, wherein detecting the presence of proximity speech in the ambient sound comprises detecting stationary background noise.

30. The method of Claim 25, wherein detecting the presence of proximity speech in the ambient sound comprises detecting speech from a close-talking proximity talker.

31. The method of Claim 25, wherein detecting the presence of proximity speech in the ambient sound comprises detecting from the at least one input signal a spectral flatness measure of the ambient sound, wherein detecting the spectral flatness measure of the ambient sound comprises detecting variation in spectral content of the ambient sound.

32. An integrated circuit for implementing at least a portion of an audio device, comprising:

an audio output configured to reproduce audio information by generating an audio output signal for communication to at least one transducer of the audio device;

a microphone input configured to receive an input signal indicative of ambient sound external to the audio device; and

a processor configured to:

detect from the input signal a near-field sound in the ambient sound; and

modify a characteristic of the audio information in response to detection of the near-field sound.

33. The integrated circuit of Claim 32, the processor further configured to:

determine from the at least one input signal a direction of the ambient sound; and modify the characteristic of the audio information reproduced to the at least one transducer in response to the direction of the ambient sound indicating that the ambient sound is sound from a user of the audio device.

34. The integrated circuit of Claim 32, the processor further configured to:

determine from the at least one input signal a direction of the ambient sound; and

modify the characteristic of the audio information reproduced to the at least one transducer in response to the direction of the ambient sound indicating that the ambient sound is speech from a user of the audio device.

35. The integrated circuit of Claim 32, wherein modifying the characteristic of the audio information comprises attenuating the audio information.

36. The integrated circuit of Claim 32, the processor further configured to modify the characteristic of the audio information in response to a detection of the near-field sound being persistent for at least a predetermined time.

37. The integrated circuit of Claim 36, the processor further configured to:

detect from the at least one input signal absence of the near-field sound in the ambient sound; and

cease modifying the characteristic of the audio information reproduced to the at least one transducer in response to the absence of the near-field sound for at least a second predetermined time.

38. The integrated circuit of Claim 36, the processor further configured to:

in addition to detecting the near-field sound, detect from the at least one input signal ambient sound other than the near-field sound in the ambient sound; and

modify the characteristic of the audio information reproduced to the at least one transducer in response to detection of the ambient sound.

39. The integrated circuit of Claim 38, the processor further configured to:

determine from the at least one input signal a direction of the ambient sound; and

modify the characteristic of the audio information reproduced to the at least one transducer in response to the direction of the ambient sound indicating that the ambient sound is sound other than the near-field sound.

40. The integrated circuit of Claim 38, the processor further configured to:

detect from the at least one input signal whether the ambient sound comprises background noise; and

modify the characteristic of the audio information reproduced to the at least one transducer in response to detection of the background noise in the ambient sound.

41. The integrated circuit of Claim 38, the processor further configured to:

detect from the at least one input signal whether the ambient sound comprises a tonal alarm; and

modify the characteristic of the audio information reproduced to the at least one transducer in response to detection of the tonal alarm in the ambient sound.

42. The integrated circuit of Claim 41, wherein detecting the tonal alarm in the ambient sound comprises:

detecting from the at least one input signal a direction of the ambient sound;

detecting from the at least one input signal a spectral flatness measure of the ambient sound;

detecting from the at least one input signal near-field spatial statistics of the ambient sound; and

detecting the tonal alarm based on the direction of the ambient sound, the presence or absence of the background noise, and the near-field spatial statistics.

43. The integrated circuit of Claim 41, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise a correlation between the first microphone signal and the second microphone signal.

44. The integrated circuit of Claim 42, wherein detecting the direction of the ambient sound comprises determining whether the direction of the ambient sound is within an acceptance angle of near-field sound.

45. The integrated circuit of Claim 42, wherein detecting the near-field spatial statistics of the ambient sound comprises detecting whether a normalized cross-correlation statistic is greater than a threshold.

46. The integrated circuit of Claim 42, wherein detecting the spectral flatness measure of the ambient sound comprises detecting whether the noise spectrum is flat in most sub-bands of the ambient sound, but not all sub-bands of the ambient sound.

47. The integrated circuit of Claim 32, wherein detecting the near-field sound in the ambient sound comprises:

detecting from the at least one input signal a direction of the ambient sound;

detecting from the at least one input signal a presence of speech in the ambient sound;

detecting from the at least one input signal near-field spatial statistics of the ambient sound; and

detecting the near-field sound based on the direction, presence or absence the of speech, and the near-spatial statistics of the ambient sound.

48. The integrated circuit of Claim 47, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise a correlation between the first microphone signal and the second microphone signal.

49. The integrated circuit of Claim 47, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise an interference-to-signal ratio associated with near-field sound.

50. The method of Claim 47, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise an inter-microphone level difference between the first microphone signal and the second microphone signal.

51. The integrated circuit of Claim 47, wherein detecting the direction of the ambient sound comprises determining whether the direction of the ambient sound is within an acceptance angle of near-field sound.

52. The integrated circuit of Claim 47, wherein detecting the near-field spatial statistics of the ambient sound comprises:

detecting whether a normalized cross-correlation statistic is greater than a first threshold;

detecting whether an interference to near-field desired signal ratio is lesser than a second threshold; and

detecting whether an inter-microphone level difference is greater than a third threshold.

53. The integrated circuit of Claim 52, wherein the second threshold is adjusted based on an estimate of background noise in the ambient sound.

54. The integrated circuit of Claim 52, wherein the second threshold is adjusted based on an estimate of background noise in the ambient sound.

55. The integrated circuit of Claim 32, the processor further configured to:

detect from the at least one input signal a direction of the ambient sound;

detect from the at least one input signal a presence of background noise in the ambient sound;

detect from the at least one input signal a presence of proximity speech in the ambient sound;

detect from the at least one input signal a volume of the ambient sound;

detect from the at least one input signal near-field spatial statistics of the ambient sound;

detect based on the direction, presence or absence of background noise, presence or absence of the speech, the volume, and the near-spatial statistics of the ambient sound, a presence of an audio event comprising a proximity sound event; and

modify the characteristic of the audio information reproduced to the at least one transducer in response to presence of the audio event.

56. The integrated circuit of Claim 32, the processor further configured to:

detect variation in spectral content of the ambient sound; and

detect based on the direction, presence or absence of background noise, presence or absence of the speech, the volume, the near-spatial statistics of the ambient sound, and the spectral content of the ambient sound, a presence of an audio event comprising a proximity sound event.

57. The integrated circuit of Claim 56, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise a correlation between the first microphone signal and the second microphone signal.

58. The integrated circuit of Claim 56, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise an interference-to-signal ratio associated with near-field sound.

59. The integrated circuit of Claim 56, wherein:

the at least one input signal comprises a first microphone signal indicative of ambient sound at a first microphone and a second microphone signal indicative of ambient sound at a second microphone; and

the near-field spatial statistics comprise an inter-microphone level difference between the first microphone signal and the second microphone signal.

60. The integrated circuit of Claim 56, wherein detecting the presence of proximity speech in the ambient sound comprises detecting stationary background noise.

61. The integrated circuit of Claim 56, wherein detecting the presence of proximity speech in the ambient sound comprises detecting speech from a close-talking proximity talker.

62. The integrated circuit of Claim 56, wherein detecting the presence of proximity speech in the ambient sound comprises detecting from the at least one input signal a spectral flatness measure of the ambient sound, wherein detecting the spectral flatness measure of the ambient sound comprises detecting variation in spectral content of the ambient sound.

63. A method for processing audio information in an audio device comprising:

reproducing audio information by generating an audio output signal for communication to at least one transducer of the audio device;

receiving at least one input signal indicative of ambient sound external to the audio device;

detecting from the at least one input signal an audio event; and

modifying a characteristic of the audio information reproduced to the at least one transducer in response to detection of the audio event being persistent for at least a predetermined time.

64. The method of Claim 63, further comprising ceasing to modify the characteristic of the audio information reproduced to the at least one transducer in response to an absence of the audio event for at least a second predetermined time.

65. The method of Claim 63, wherein the audio event comprises at least one of a near-field event, a proximity event, and an alarm event.

66. An integrated circuit for implementing at least a portion of an audio device, comprising:

an audio output configured to reproduce audio information by generating an audio output signal for communication to at least one transducer of the audio device;

a microphone input configured to receive an input signal indicative of ambient sound external to the audio device; and

a processor configured to:

detect from the input signal an audio event; and

modify a characteristic of the audio information reproduced to the at least one transducer in response to detection of the audio event being persistent for at least a predetermined time.

67. The integrated circuit of Claim 66, the processor further configured to cease modifying the characteristic of the audio information reproduced to the at least one transducer in response to an absence of the audio event for at least a second predetermined time.

68. The integrated circuit of Claim 66, wherein the audio event comprises at least one of a near-field event, a proximity event, and an alarm event.

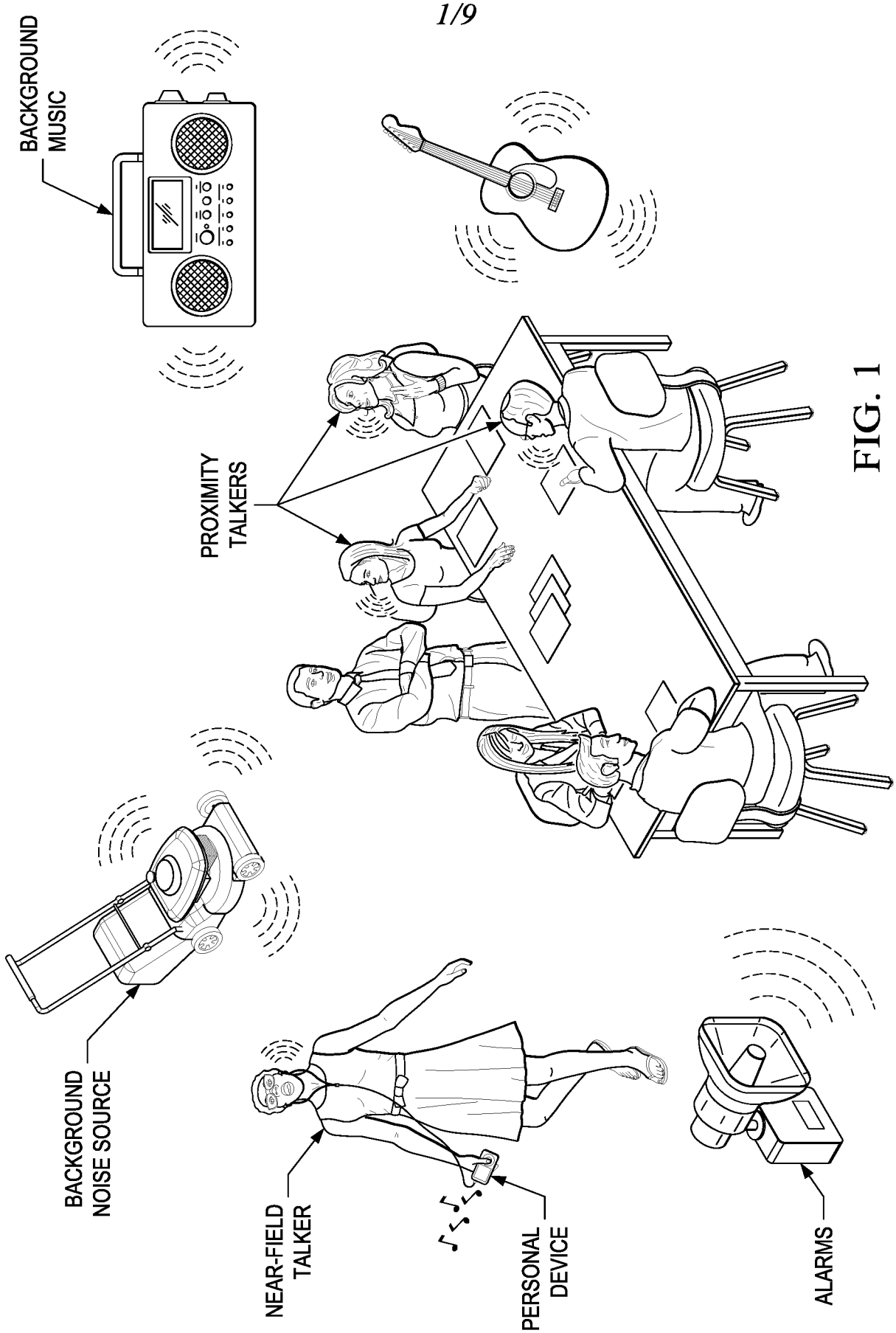

BACKGROUND MUSIC
PROXIMITY TALKERS
BACKGROUND NOISE SOURCE
NEAR-FIELD TALKER
PERSONAL DEVICE
ALARMS

FIG. 1

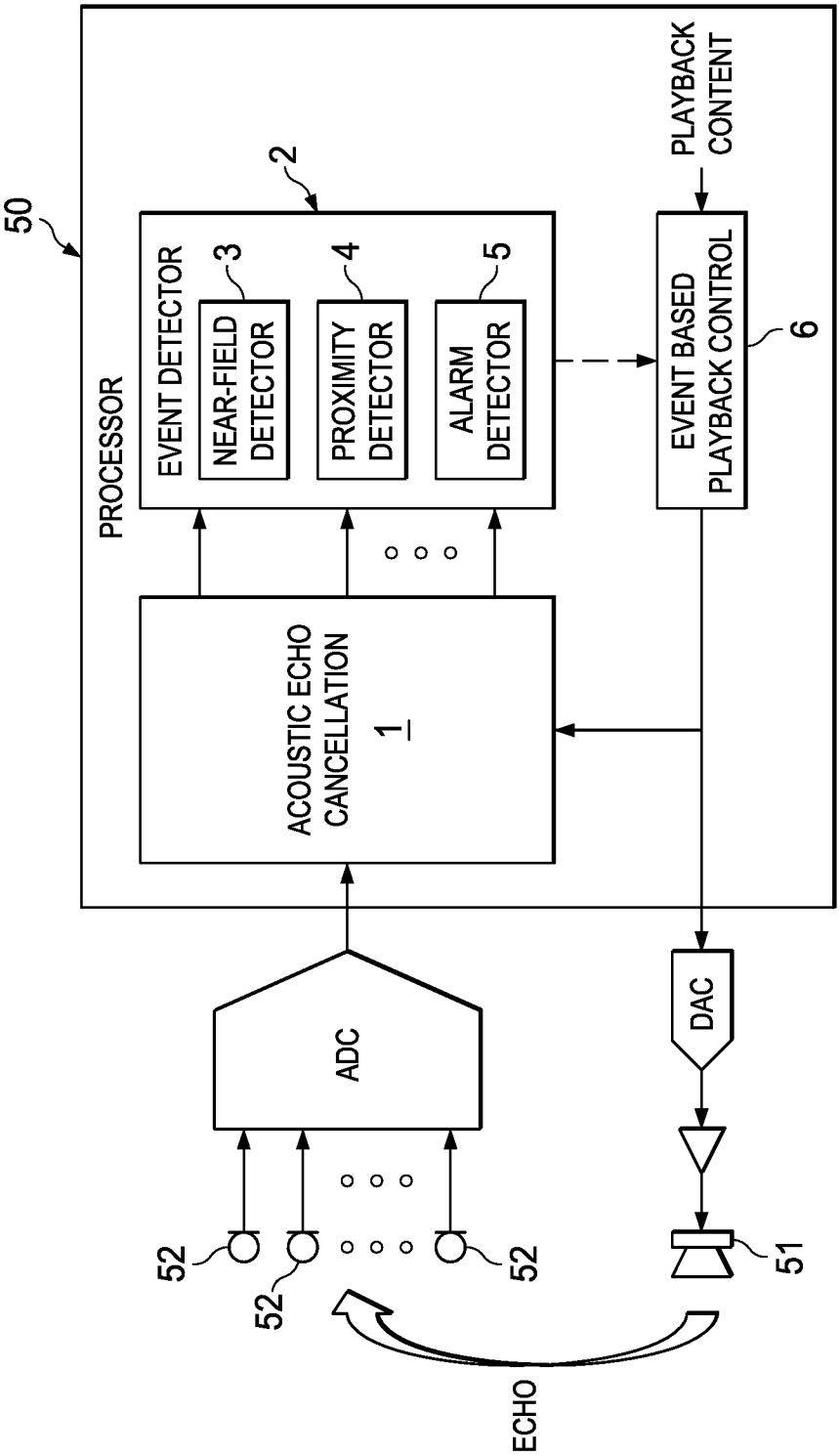

50
PROCESSOR
2
EVENT DETECTOR
3
NEAR-FIELD DETECTOR
4
PROXIMITY DETECTOR
5
ALARM DETECTOR
PLAYBACK CONTENT
EVENT BASED PLAYBACK CONTROL
6
ACOUSTIC ECHO CANCELLATION
1
ADC
DAC
52
52
52
51
ECHO

FIG. 2

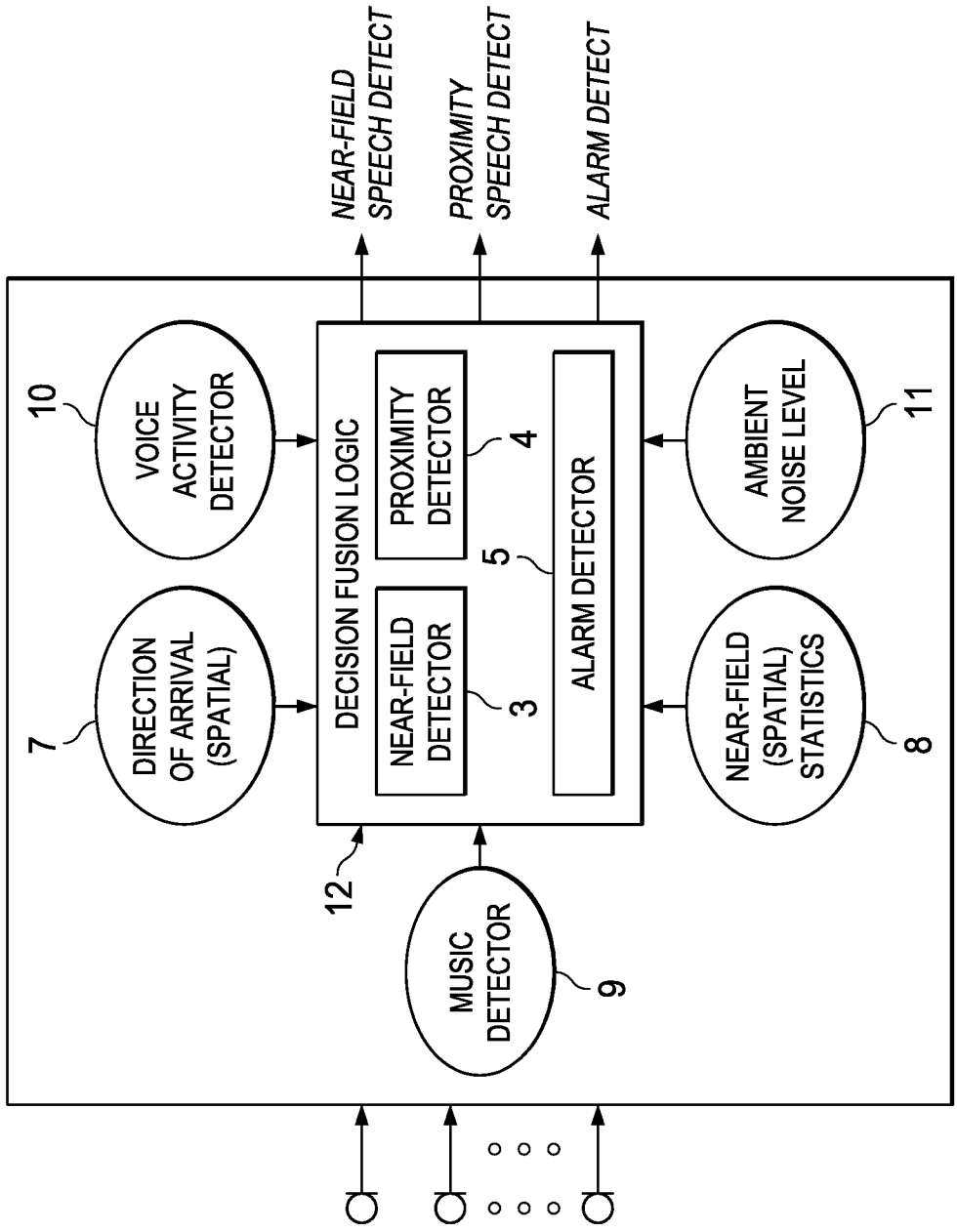

NEAR-FIELD SPEECH DETECT
PROXIMITY SPEECH DETECT
ALARM DETECT
10
VOICE ACTIVITY DETECTOR
7
DIRECTION OF ARRIVAL (SPATIAL)
DECISION FUSION LOGIC
PROXIMITY DETECTOR
4
NEAR-FIELD DETECTOR
3
5
ALARM DETECTOR
12
MUSIC DETECTOR
9
AMBIENT NOISE LEVEL
11
NEAR-FIELD (SPATIAL) STATISTICS
8

FIG. 3

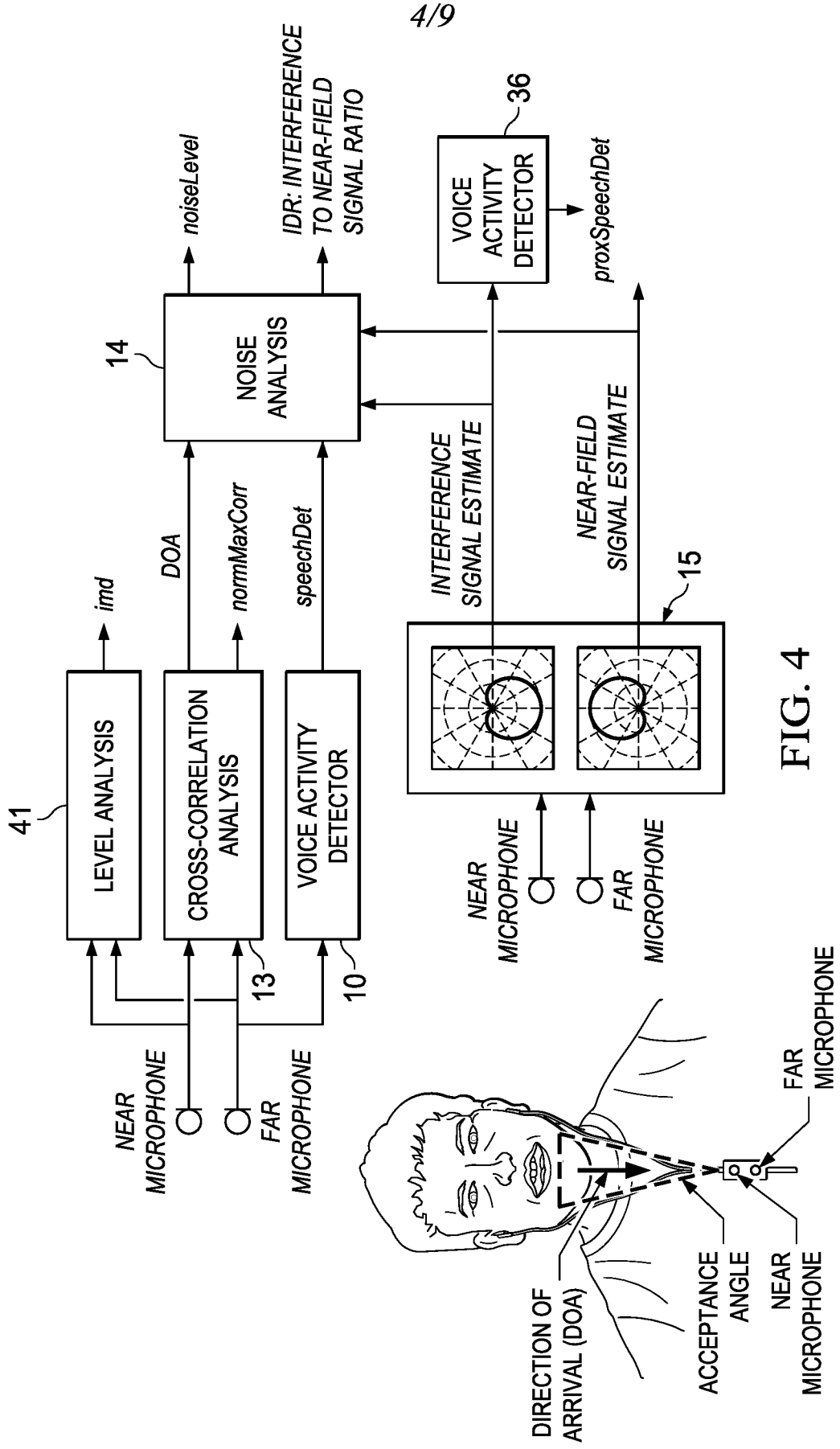

41
LEVEL ANALYSIS
imd
CROSS-CORRELATION ANALYSIS
DOA
normMaxCorr
13
VOICE ACTIVITY DETECTOR
speechDet
10
NEAR MICROPHONE
FAR MICROPHONE
14
NOISE ANALYSIS
noiseLevel
IDR: INTERFERENCE TO NEAR-FIELD SIGNAL RATIO
36
VOICE ACTIVITY DETECTOR
proxSpeechDet
INTERFERENCE SIGNAL ESTIMATE
NEAR-FIELD SIGNAL ESTIMATE
15
NEAR MICROPHONE
FAR MICROPHONE
DIRECTION OF ARRIVAL (DOA)
ACCEPTANCE ANGLE
NEAR MICROPHONE
FAR MICROPHONE

FIG. 4

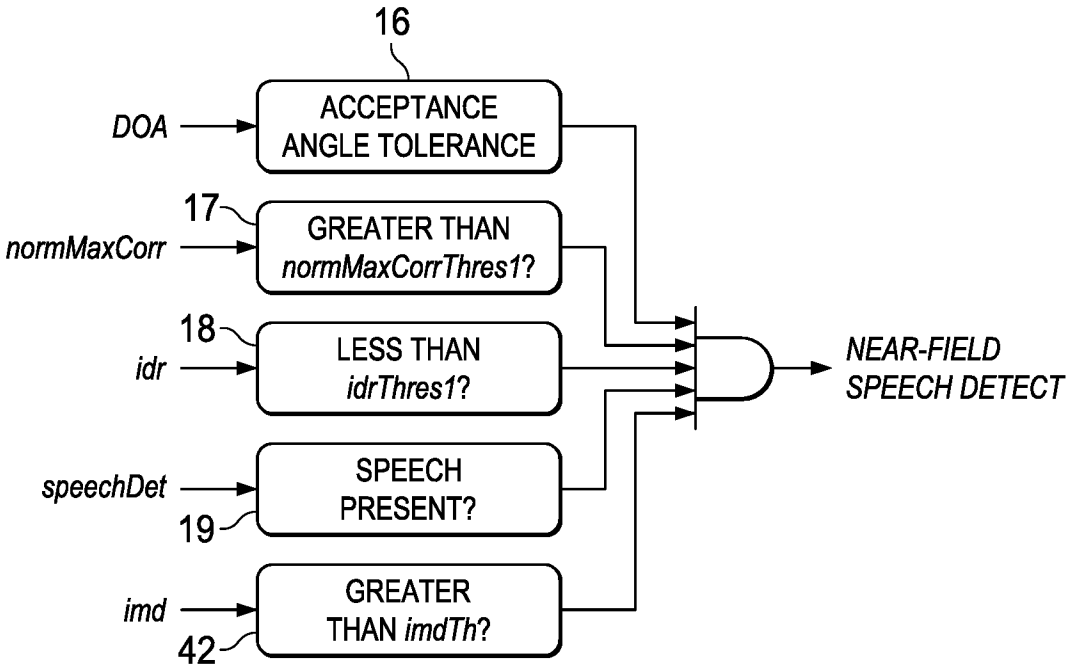
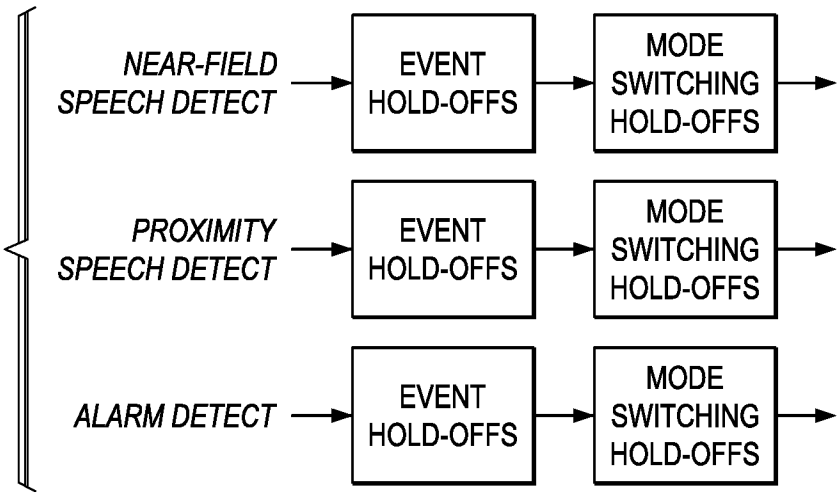

16
DOA
ACCEPTANCE ANGLE TOLERANCE
17
normMaxCorr
GREATER THAN normMaxCorrThres1?
18
idr
LESS THAN idrThres1?
NEAR-FIELD SPEECH DETECT
speechDet
SPEECH PRESENT?
19
imd
GREATER THAN imdTh?
42

FIG. 5

FIG. 10


NEAR-FIELD SPEECH DETECT
EVENT HOLD-OFFS
MODE SWITCHING HOLD-OFFS
PROXIMITY SPEECH DETECT
EVENT HOLD-OFFS
MODE SWITCHING HOLD-OFFS
ALARM DETECT
EVENT HOLD-OFFS
MODE SWITCHING HOLD-OFFS

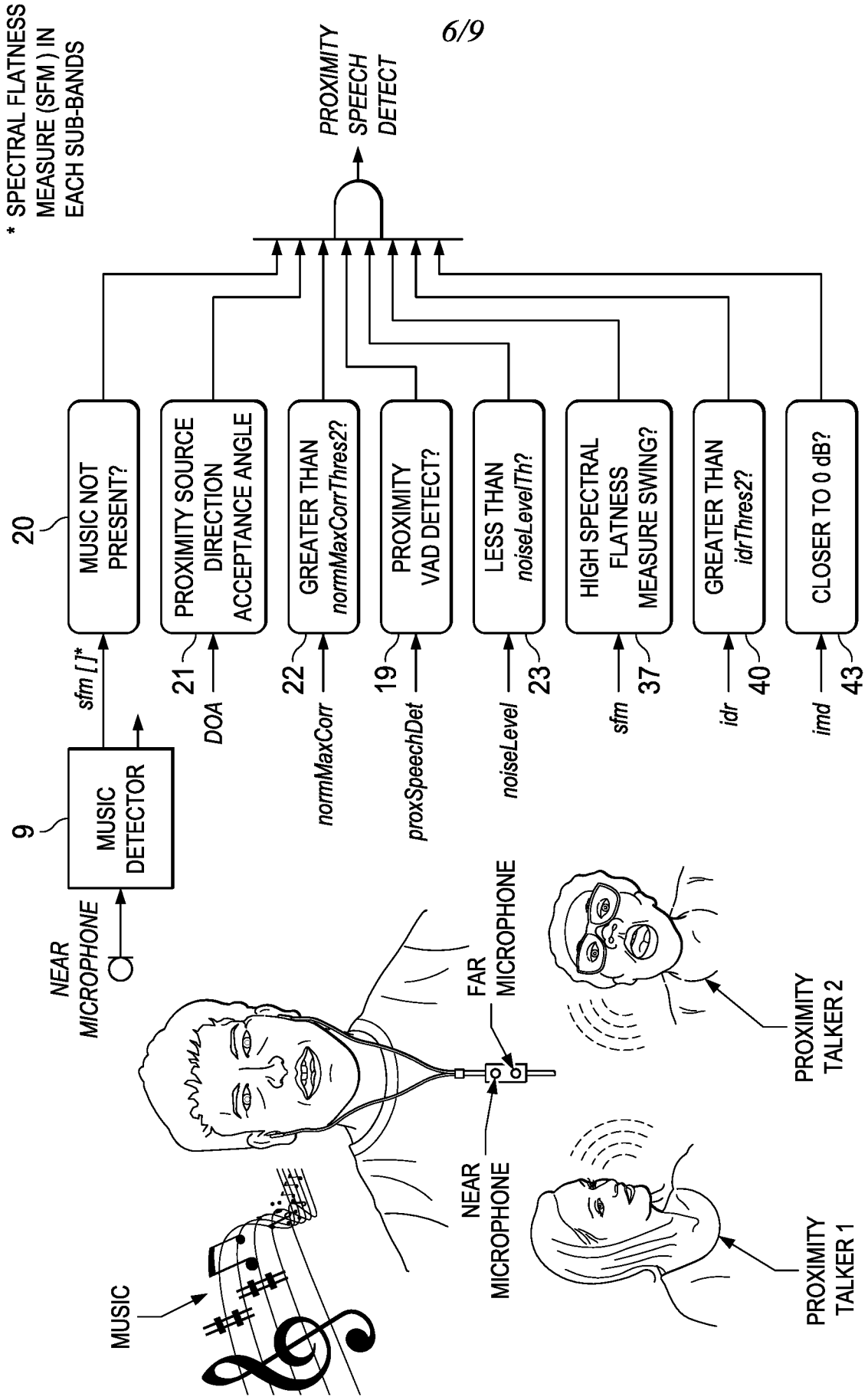
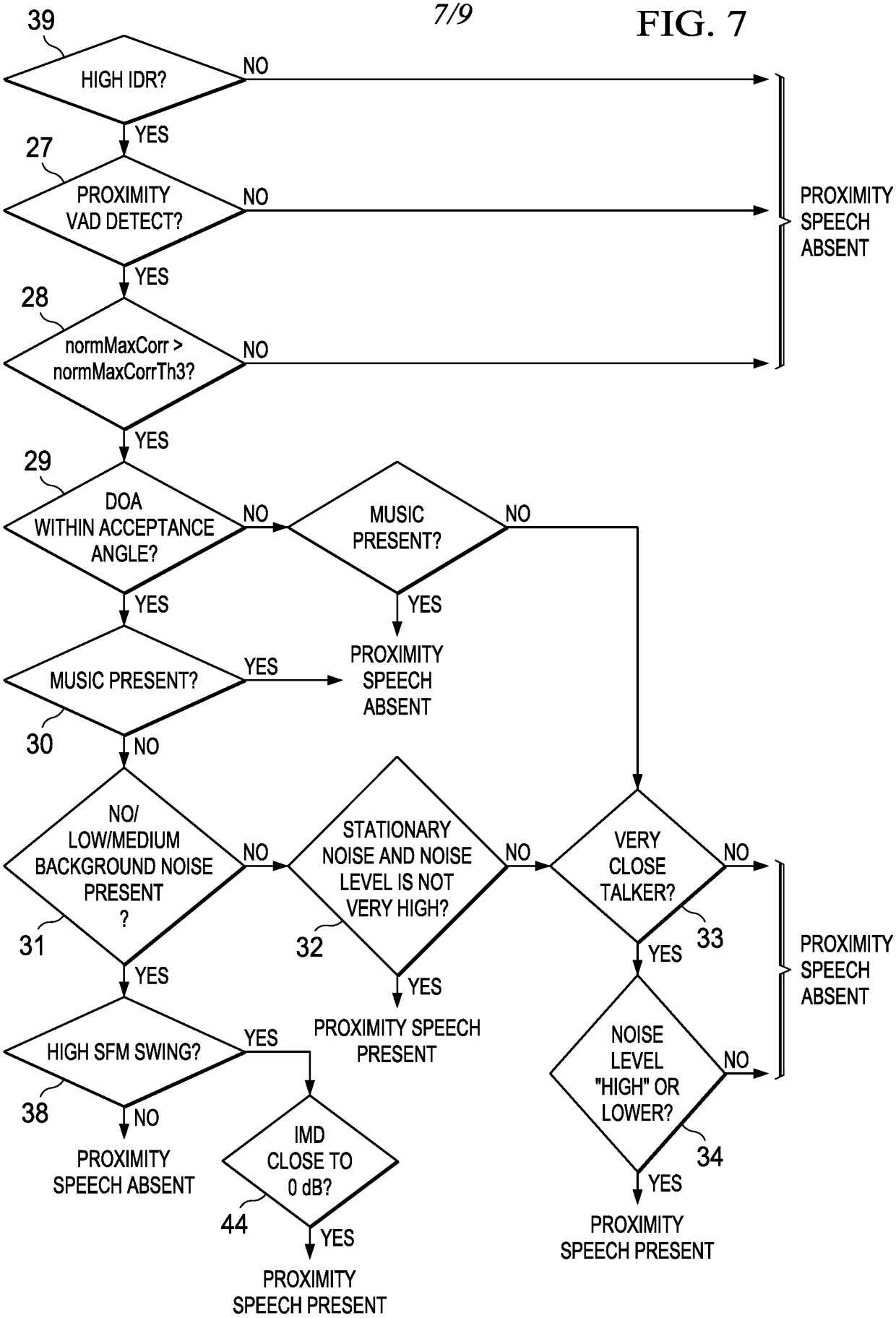

* SPECTRAL FLATNESS MEASURE (SFM) IN EACH SUB-BANDS
PROXIMITY SPEECH DETECT
20
MUSIC NOT PRESENT?
PROXIMITY SOURCE DIRECTION ACCEPTANCE ANGLE
GREATER THAN normMaxCorrThres2?
PROXIMITY VAD DETECT?
LESS THAN noiseLevelTh?
HIGH SPECTRAL FLATNESS MEASURE SWING?
GREATER THAN idrThres2?
CLOSER TO 0 dB?
sfm [ ]*
21
DOA
22
normMaxCorr
19
proxSpeechDet
23
noiseLevel
37
sfm
40
idr
43
imd
9
MUSIC DETECTOR
NEAR MICROPHONE
FAR MICROPHONE
NEAR MICROPHONE
MUSIC
PROXIMITY TALKER 1
PROXIMITY TALKER 2

FIG. 6

FIG. 7


39
HIGH IDR?
NO
YES
27
PROXIMITY VAD DETECT?
NO
YES
28
normMaxCorr > normMaxCorrTh3?
NO
YES
PROXIMITY SPEECH ABSENT
29
DOA WITHIN ACCEPTANCE ANGLE?
NO
MUSIC PRESENT?
NO
YES
YES
MUSIC PRESENT?
YES
PROXIMITY SPEECH ABSENT
30
NO
NO/ LOW/MEDIUM BACKGROUND NOISE PRESENT ?
NO
STATIONARY NOISE AND NOISE LEVEL IS NOT VERY HIGH?
NO
VERY CLOSE TALKER?
NO
31
32
33
YES
YES
YES
PROXIMITY SPEECH ABSENT
HIGH SFM SWING?
YES
PROXIMITY SPEECH PRESENT
NOISE LEVEL "HIGH" OR LOWER?
NO
38
NO
IMD CLOSE TO 0 dB?
34
YES
PROXIMITY SPEECH ABSENT
44
YES
YES
PROXIMITY SPEECH PRESENT
PROXIMITY SPEECH PRESENT

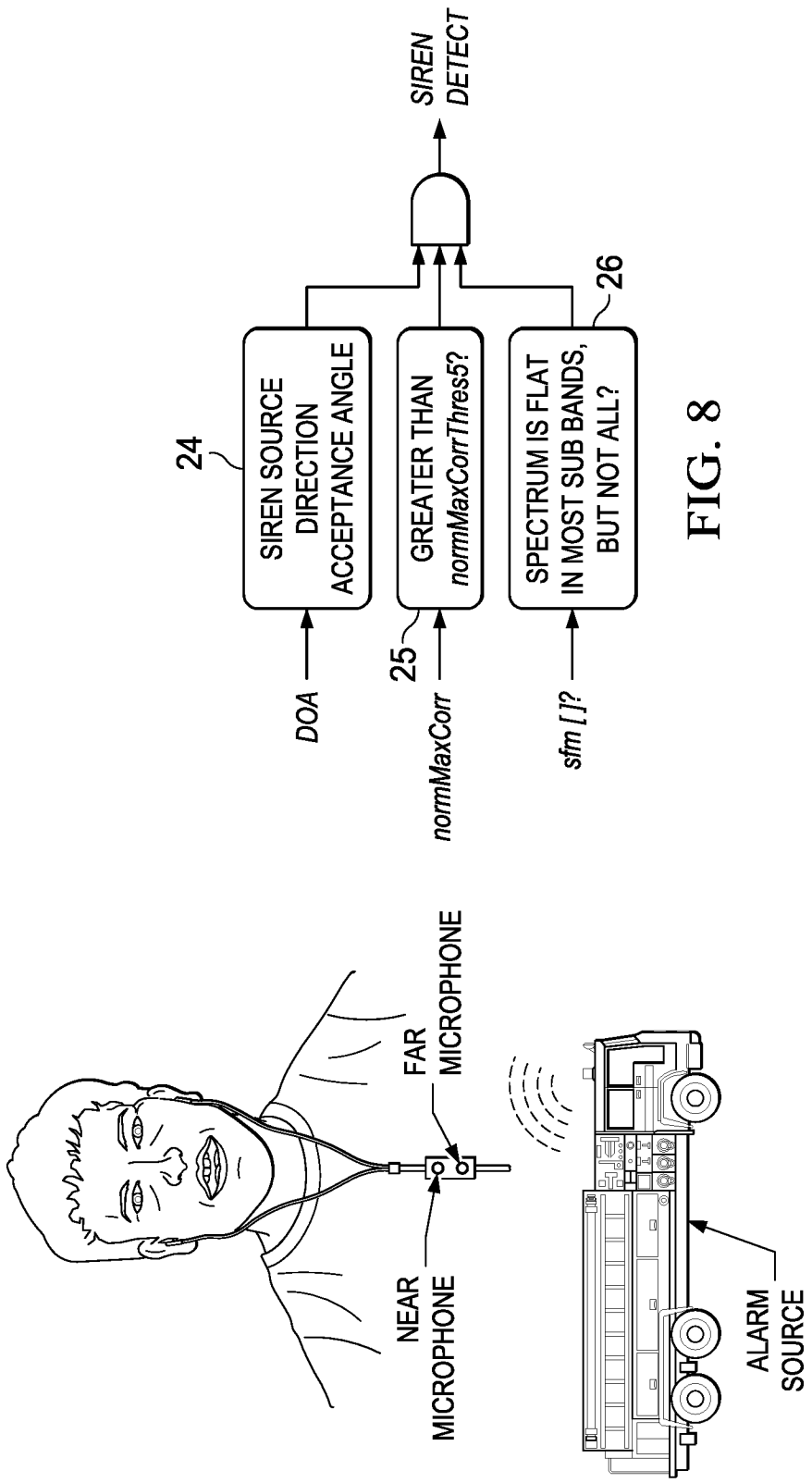

SIREN DETECT
24
SIREN SOURCE DIRECTION ACCEPTANCE ANGLE
GREATER THAN normMaxCorrThres5?
SPECTRUM IS FLAT IN MOST SUB BANDS, BUT NOT ALL?
26
25
DOA
normMaxCorr
sfm [ ]?
FAR MICROPHONE
NEAR MICROPHONE
ALARM SOURCE

FIG. 8

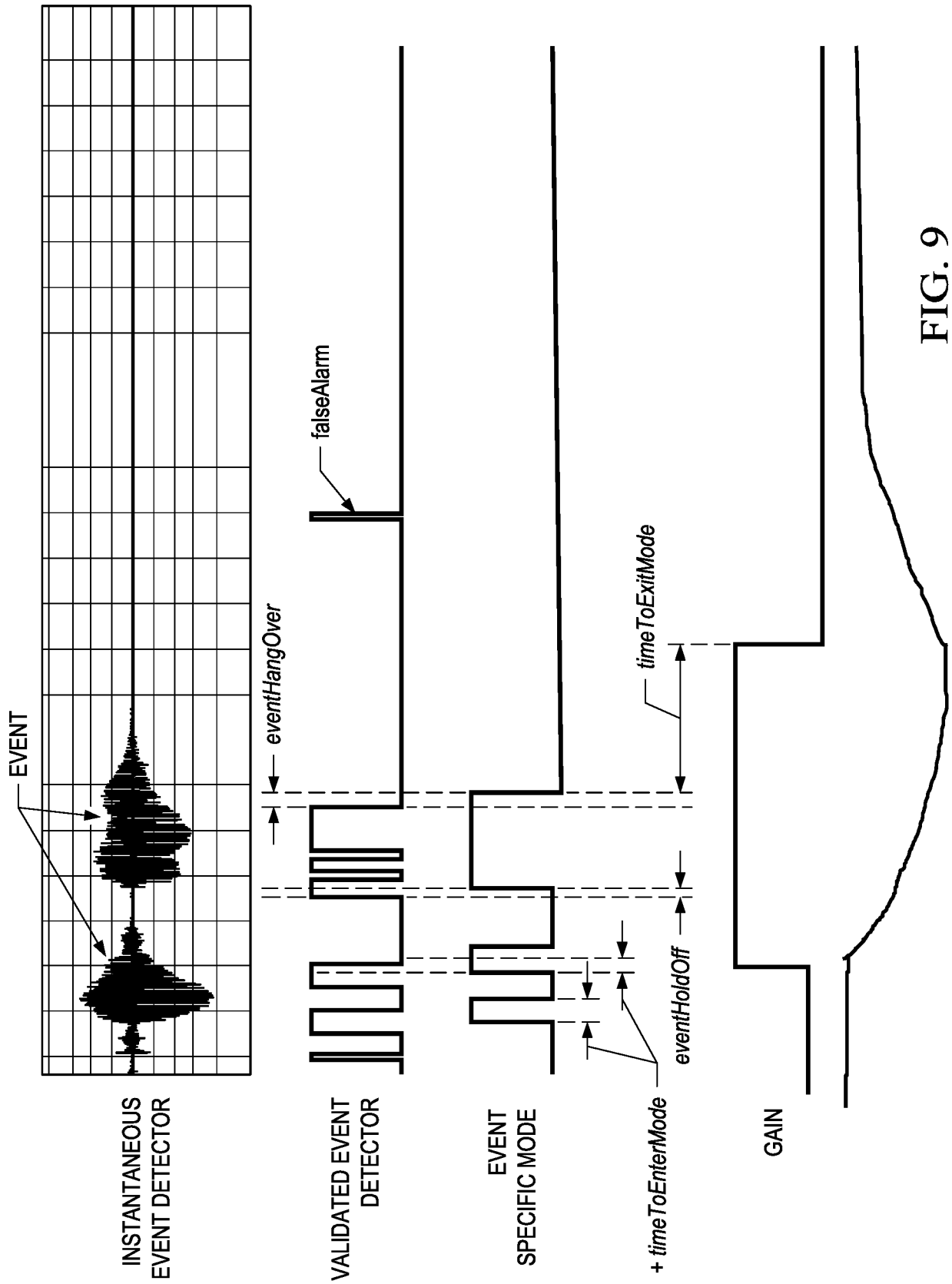

EVENT
INSTANTANEOUS EVENT DETECTOR
eventHangOver
falseAlarm
VALIDATED EVENT DETECTOR
EVENT SPECIFIC MODE
timeToExitMode
+ timeToEnterMode
eventHoldOff
GAIN

FIG. 9