(72) Inventors:
• Schmidt, Gerhard Uwe
89081 Ulm (DE)
• Buck, Markus
88400 Biberach (DE)
• Haulick, Tim
89143 Blaubeuren (DE)

(74) Representative: Grünecker, Kinkeldey,
Stockmair & Schwanhäusser
Anwaltssozietät
Maximilianstrasse 58
80538 München (DE)

(54) **Automatic recognition of vehicle operation noises**

(57) The present invention relates to a system for automatic recognition of operation noises of a vehicle, comprising at least one microphone installed in a vehicular cabin for detecting acoustic signals and generating microphone signals, a database comprising speech templates and operation noise templates, feature extracting means configured to receive the generated microphone signals and to extract at least one set of noise feature parameters and/or at least one set of speech feature parameters from the generated microphone signals, a speech and noise recognition means configured to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters and/or to determine at least one speech template that best matches the at least one extracted set of speech feature parameters and a control means configured to control the speech and noise recognition means to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters and/or to determine at least one speech template that best matches the at least one extracted set of speech feature parameters. The invention also relates to a method for recognizing operation noises of a vehicle comprising providing a speech recognition system comprising a database comprising speech templates and operation noise templates, extracting at least one set of noise feature parameters and/or at least one set of speech feature parameters from microphone signals generated from acoustic signals and determining at least one operation noise template that best matches the at least one extracted set of noise feature parameters and/or determining at least one speech template that best matches the at least one extracted set of speech feature parameters.
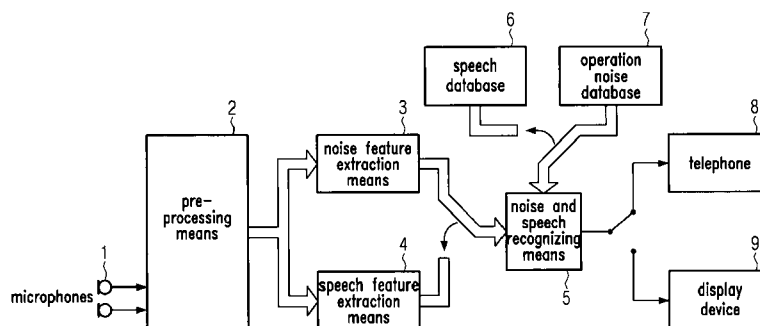
FIG. 1

EP 1 703 471 A1

Description

Field of Invention

[0001]   The present invention relates to the diagnosis of vehicle operation and, in particular, to the automatic recognition of vehicle operation noises by means of microphones to detect present or future operation faults.

Prior Art

[0002]   The diagnosis of the operation of vehicles is an important task in order to prevent severe failures and to improve the overall safety of the passengers. In recent years, automobiles have been equipped with a variety of electronic diagnosis devices that are able to permanently sample data that may be helpful for the personnel of service stations in detecting faults during routine inspections and in determining the cause for actually occurred failures. Additionally, oscilloscopes are commonly used in service stations to measure and monitor signals generated by electronic and electrical components.

[0003]   Remote vehicle diagnosis allows for wirelessly transmitting data sampled by vehicle sensors to databases of service stations. Thus, immediate support is made available. Drivers may even receive warnings from service stations in case of the remote detection of severe failures of the vehicle operation.

[0004]   Acoustic signals represent an important information source for the state of operation of a vehicle, in particular, of the engine and operatively connected components. Usually, skilled motorcar mechanics are able to guess or even determine failures when listening to operation noises.

[0005]   However, the common driver is not able to use the acoustic information for diagnosis purposes. In addition, the hearing of most of the drivers shows only a limited frequency range. Moreover, some creeping evolution of a malfunction might scarcely be detectable, since the associated acoustic variations are hardly ever perceptible.

[0006]   Present vehicle diagnosis systems including audio analysis means require sensors installed outside the vehicular cabin for the monitored components. Such sensors show their own faults, in particular, when aging and suffer, e.g., from corrosion.

[0007]   Consequently, there is still a need for a more comfortable and reliable audio diagnosis of a vehicle operation that, in particular, is not hampered by the expensive employment of multiple sensors showing only limited reliability.

Description of the invention

[0008]   The above mentioned object is achieved by a system for automatic recognition of operation noises of a vehicle according to claim 1 and a method for recognizing operation noises of a vehicle according to claim 16.

[0009]   According to claim 1 it is provided a system for automatic recognition of operation noises of a vehicle, comprising at least one microphone installed in a vehicular cabin for detecting acoustic signals and generating microphone signals;
a database comprising speech templates and operation noise templates;
feature extracting means configured to receive the generated microphone signals and to extract at least one set of noise feature parameters and/or at least one set of speech feature parameters from the generated microphone signals;
a speech and noise recognition means configured to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters and/or to determine at least one speech template that best matches the at least one extracted set of speech feature parameters; and
a control means configured to control the speech and noise recognition means to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters and/or
to determine at least one speech template that best matches the at least one extracted set of speech feature parameters.

[0010]   Recognition of operation noises comprises classifying and/or identifying these noises. Classes of operation noises can comprise, e.g., wheel bearing noise, ignition noise, braking noise, engine noise depending on the engine speed etc., and each class may comprise sub-classes for noise samples representing, e.g., regular, critical and super-critical operation noise levels and frequency ranges. Both the noise and the speech templates represent trained/learned model samples of particular acoustic signals and advantageously comprise feature (characteristic) vectors for the particular acoustic signals comprising relevant feature parameters as, e.g., the cepstral coefficients or amplitudes per frequency bin.

[0011]   The training is preferably carried out in collaboration with skilled mechanics and by detecting and recording the operation noises of vehicles showing commonly occurring faults and of vehicles that ideally operate faultlessly. It may be advantageous to carry out training specific for each vehicle model. Such an individual training and generation of operation noise templates is relatively time-consuming, but enhances the reliability of the noise recognition.

[0012]   At least one microphone is used to detect acoustic signals and to generate microphone signals. It may be preferred to use more than one microphone and, in particular, at least one microphone array. Moreover, more than one

microphone array may advantageously be employed.

[0013] The microphone signals may be pre-processed, in particular, discretized and quantized, by a Fourier transformation before being input in the feature extracting means. The feature extracting means is configured to extract predetermined feature parameters from the pre-processed microphone signals, i.e. a set of feature parameters comprising at least one feature vector containing feature parameters, is generated corresponding to the acoustic signals. Such vectors may comprise about 10 to 20 feature parameters and may be calculated every 10 or 20 msec, e.g., from short-term power spectra for multiple subbands.

[0014] Noise signals within acoustic signals are assigned to one or more best matching noise templates of a database. Specifically, the feature vectors comprising feature parameters and generated by the feature extraction means may be compared with feature vectors representing said operation noise templates. These noise templates may comprise previously generated templates and also templates calculated, e.g., by some averaging, from previously generated noise templates.

[0015] Generation of the noise templates may be performed by detecting noise caused by the regular operation and different kinds of faulty operation of vehicle components. Noise templates that represent noise associated with some technical failures may be considered as elements of a particular set of fault-indicating templates.

[0016] Typical feature parameters for speech signals are, e.g., amplitudes, cepstral coefficients and predictor coefficients. Noise feature parameters may include some of the speech feature parameters or appropriate modifications thereof as highly resolved bandpass power levels in the low-frequency range.

[0017] Due to the inventive assignment of noise signals within detected acoustic signals to best matching noise templates of a database making use of the noise feature parameters, a comfortable and reliable audio diagnosis device for detecting and monitoring a vehicle operation is provided by the invention. Surprisingly, speech recognition system that become increasingly prevalent in vehicular cabins can rather readily be modified, mainly on a software basis, to be usable for the disclosed diagnosis of vehicle operation based on acoustic signals. Tools known from speech recognition can widely be adapted and the skilled person can easily incorporate modifications useful for the classification of noise signals. Apparently, the synergetic effects are rather significant.

[0018] It may be noted that, whereas the present invention is regarded as being particularly useful for automobiles, different vehicles, as watercrafts and aircrafts, may also be included in the term 'vehicle' as used herein.

[0019] Employment of a control means is an important feature of the present invention. The detected acoustic signals and the generated microphone signals may comprise speech as well as noise information. For reasons of, e.g., limited computer resources as limited memory and CPU power, it may be preferred not to perform both the speech recognition and noise recognition processes in parallel.

[0020] If, e.g., a passenger of the vehicle wants explicitly to use the speech recognition means, noise recognition may be stopped or disabled, in order to have the entire computing power available for the speech recognition processing. If, on the other hand, a passenger switches off the speech recognition operation, noise recognition may be performed exclusively, i.e., in particular, at least one operation noise template that best matches the at least one extracted set of noise feature parameters can be determined.

[0021] The control means may be configured to control the feature extracting means to extract at least one set of noise feature parameters, if it controls the speech and noise recognition means to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters, and to extract at least one set of speech feature parameters, if it controls the speech and noise recognition means to determine at least one speech template that best matches the at least one extracted set of speech feature parameters. Thereby, the computer resources are managed even more effectively.

[0022] The control means can be configured to control the speech and noise recognition means to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters, if the acoustic signals do not comprise speech signals for at least a predetermined time period.

[0023] It may be determined, e.g., by the feature extracting means, that the acoustic signals do not contain any speech signals. In this case no speech analysis and processing is necessary and accordingly it may be advantageous to safe all computing power for the noise recognition. The predetermined time period may be manually set by a user.

[0024] According to an embodiment of the inventive system, a push-to-talk lever may further be provided and in this case the control means may be configured to control the speech and noise recognition means to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters, if the push-to-talk lever is pushed in an "off"-position and/or to control the speech and noise recognition means to determine at least one speech template that best matches the at least one extracted set of speech feature parameters, if the push-to-talk lever is pushed in an "on"-position.

[0025] Accordingly, a user, e.g., the driver, can manually choose from noise and speech recognition performed by the system. Reliability and ease of use can thus, be improved.

[0026] Preferably, the system for automatic recognition of operation noises of a vehicle may further comprise at least one application means configured to perform applications on the basis of the at least one determined best matching

speech template or the at least one determined best matching operation noise template.

[0027] If, e.g., a speech template representing a phone number is identified, this number may be dialed by a mobile phone representing an application means that is connected to the noise and speech recognition means. If the at least one application means comprises a display, information corresponding to an identified operation noise template may be shown on the display.

[0028] The at least one application means may comprise a warning means configured to output an acoustic and/or visual and/or haptic warning, if the speech and noise recognition means is controlled to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters and if the difference between the extracted noise feature parameters and the noise feature parameters of the operation noise template determined to best match the at least one extracted set of noise feature parameters exceeds a predetermined level or if the operation noise template determined to best match the at least one extracted set of noise feature parameters is an element of a predetermined set of particular operation noise templates indicative for operation faults.

[0029] The difference between the extracted noise feature parameters and the noise feature parameters of the operation noise template can be measured by an appropriate distance measure as commonly used in the art. The predetermined level can be set during a training phase. Operation noise templates indicative for operation faults are usually trained before installation of the system in a vehicle.

[0030] Thus, a driver of the vehicle may be warned, if some failure actually affects the operation of the vehicle or is to be expected to affect faultless operation in the near future. The driver can react accordingly and avoid severe damages and risks.

[0031] The at least one application means may also comprise a wireless communication device configured to transmit, in particular, to a service center, the best matching operation noise template and/or the at least one extracted set of noise feature parameters and/or the generated microphone signals. The wireless communication device may be a mobile phone.

[0032] On the basis of the received data skilled mechanics may be informed about the operation and safety status of a vehicle and may warn and support the driver in case of severe failures by telecommunication.

[0033] The wireless communication device may be configured to automatically transmit data comprising the best matching operation noise template and/or the at least one extracted set of noise feature parameters and/or the generated microphone signals, if the difference between the extracted noise feature parameters and the noise feature parameters of the operation noise template determined to best match the at least one extracted set of noise feature parameters exceeds a predetermined level and/or if the operation noise template determined to best match the at least one extracted set of noise feature parameters is an element of a predetermined set of particular operation noise templates indicative for operation faults.

[0034] The automatic transmission of data comprising information about the operation noises and thereby the operation state of the vehicle improves safety and comfort.

[0035] The at least one application means may comprise a speech output configured to output a verbal warning, if the difference between the extracted noise feature parameters and the noise feature parameters of the operation noise template determined to best match the at least one extracted set of noise feature parameters exceeds a predetermined level and/or if the operation noise template determined to best match the at least one extracted set of noise feature parameters is an element of a predetermined set of particular operation noise templates indicative for operation faults.

[0036] The driver my even be given detailed instructions how to react on a given failure or expected failure in the operation of the vehicle. Thereby, safety and ease of use can further be increased by a synthesized speech output.

[0037] According to one embodiment the system for automatic recognition of operation noises of a vehicle may further comprise at least one vehicle component sensor configured to generate sensor signals and the speech and noise recognition means may be configured to determine the at least one operation noise template that best matches the at least one extracted set of noise feature parameters partly on the basis of the generated sensor signals.

[0038] Information by vehicle component sensors known in the art, as e.g., sensors for the engine speed, may assist the speech and noise recognition means in determining the best matching operation noise template, e.g., by reducing the set of the possible candidate templates.

[0039] If the speech and recognition means is provided with signals containing information about the engine speed, e.g., the reliability of the recognizing result may be improved. Moreover, the operation of application means may be influenced by sensor data. For example, one of the application means may be a device to reduce the engine speed in cases of very severe faults identified by the system for recognition of operating noises.

[0040] Sensor signals may be synchronized with the microphone signals and the noise and speech recognizing means may make use of both, the sensor signals and the microphone signals, to improve performance of the recognizing process.

[0041] As mentioned above the microphone signals may be generated by one or more microphone arrays. A microphone array may comprise at least one first microphone configured for usage in common speech recognition systems and/or speech dialog systems and/or vehicle hands-free sets and/or at least one second microphone capable of detecting acoustic signals with frequencies below and/or above the frequency range detected by the at least one first microphone.

[0042] If only microphones are used that are employed in existing speech dialog systems or speech recognition systems, almost no hardware modifications are necessary to install the disclosed system for recognition of operation noises in vehicles that are equipped with such speech processing devices.

[0043] Whereas employment of already installed microphones for detecting speech signals is advantageous in respect of costs reduction, it may be preferred to install additional microphones that are able to detect, e.g., frequency ranges below and/or above the frequencies covered by verbal utterances. Usage of microphones specially designed for frequency ranges above and, in particular, below the frequency range detected by the microphones commonly installed in vehicular cabins may significantly improve the noise recognition.

[0044] Furthermore, the at least one microphone array that can advantageously be employed can comprise at least one directional microphone, in particular, more than one directional microphone pointing in different directions, thereby improving the reliability of the recognition process and also providing a better possibility for the localization of possibly detected operations faults. If, e.g., a wheel bearing fault is detected, employment of directional microphones may be helpful in determining which one of the typically four wheel bearings shows the fault.

[0045] Moreover, the microphone signals may be beamformed by a beamforming means, in particular, an adaptive beamforming means. This action can be implemented not only to enhance the intelligibility of speech but also to improve the quality of noise signals in order to improve the reliability of the identification of the associate stored noise template. The beamformed microphone signals may be further prep-processed and eventually input in the feature extracting means.

[0046] One may also employ an inversely operating beamforming means that synchronizes microphone signals including operation noise and outputs beamformed signals with an enhance noise-to-signal level for an improved noise recognition. In that case, spatial nulls can be placed (fixed or adaptively) in the direction of the passengers in order to suppress speech signals while maintaining noise components.

[0047] Furthermore, an embodiment of the disclosed system may comprise a recording means for recording the best matching operation noise template and/or the at least one extracted set of noise feature parameters and/or the microphone signals. The recorded data can, e.g., subsequently be used for further analysis during inspection in a service station.

[0048] The present invention also provides a method for recognizing operation noises of a vehicle comprising the steps of

providing a speech recognition system comprising a database comprising speech templates and operation noise templates;

extracting at least one set of noise feature parameters and/or at least one set of speech feature parameters from microphone signals generated from acoustic signals by at least one microphone installed in a vehicular cabin; and

determining at least one operation noise template that best matches the at least one extracted set of noise feature parameters and/or determining at least one speech template that best matches the at least one extracted set of speech feature parameters.

[0049] In principle, speech and noise recognition may be performed in parallel, but it may be preferred, e.g., to safe computer resources, to determine alternatively the best matching noise template or the best matching speech template.

[0050] According to an embodiment of the method at least one set of noise feature parameters may be extracted and at least one operation noise template that best matches the at least one extracted set of noise feature parameters may be determined, if the acoustic signals do not comprise speech signals for at least a predetermined time period as it may be determined by a feature extracting means that is suitable to extract sets of noise feature parameters and speech feature parameters.

[0051] In another embodiment of the method at least one set of noise feature parameters is extracted and at least one operation noise template that best matches the at least one extracted set of noise feature parameters is determined, if a push-to-talk lever is pushed in an "off"-position and at least one set of speech feature parameters is extracted and at least one speech template that best matches the at least one extracted set of speech feature parameters is determined, if a push-to-talk lever is pushed in an "on"-position.

[0052] Moreover, the method may comprise the step of outputting an acoustic and/or visual and/or haptic warning, if the difference between the extracted noise feature parameters and the noise feature parameters of the operation noise template determined to best match the at least one extracted set of noise feature parameters exceeds a predetermined level or if the operation noise template determined to best match the at least one extracted set of noise feature parameters is an element of a predetermined set of particular operation noise templates indicative for operation faults.

[0053] The method may include transmitting of the best matching operation noise template and/or the at least one extracted set of noise feature parameters and/or the generated microphone signals by a wireless communication device, in particular, to a service station.

Transmission may be performed automatically or on a demand by a user, e.g., the driver of the vehicle.

[0054] If a wireless communication device is provided, the microphone signals may automatically be transmitted, if the difference between the extracted noise feature parameters and the noise feature parameters of the operation noise template determined to best match the at least one extracted set of noise feature parameters exceeds a predetermined level or if the operation noise template determined to best match the at least one extracted set of noise feature parameters

is an element of a predetermined set of particular operation noise templates indicative for operation faults.

[0055] The method may comprise outputting of a verbal warning, if the difference between the extracted noise feature parameters and the noise feature parameters of the operation noise template determined to best match the at least one extracted set of noise feature parameters exceeds a predetermined level or if the operation noise template determined to best match the at least one extracted set of noise feature parameters is an element of a predetermined set of operation noise templates indicative for operation faults.

[0056] Moreover, the best matching operation noise template and/or the at least one extracted set of noise feature parameters and/or the microphone signals can be stored for a subsequent analysis.

[0057] In an embodiment of the method at least one vehicle component sensor configured to generate sensor signals may be provided and in this case the determining of the at least one operation noise template that best matches the at least one extracted set of noise feature parameters can be partly based on the sensor signals.

[0058] The microphone signals used in the method for recognizing operation noises of a vehicle can be generated by at least one first microphone configured for usage in common speech recognition systems and/or speech dialog systems and/or vehicle hands-free sets and/or at least one second microphone capable of detecting acoustic signals with frequencies below and/or above the frequency range detected by the at least one first microphone.

[0059] In particular, the microphone signals can be generated by at least one directional microphone, in particular, more than one directional microphone pointing in different directions and moreover, the microphone signals may advantageously be beamformed, in particular, by an adaptive beamforming means, before at least one set of noise feature parameters and/or at least one set of speech feature parameters are extracted from the microphone signals.

[0060] Furthermore, the present invention provides a computer program product, comprising one or more computer readable media having computer-executable instructions for performing the steps of embodiments of the inventive method for automatic recognition of operation noises of vehicles as described above.

[0061] Additional features and advantages of the invention will be described with reference to the drawings:

Figure 1 shows components of an example for the system for recognition of operation noises of a vehicle comprising noise and speech feature extraction means, noise and speech recognizing means, operation noise and speech database, a telephone and a display device.

Figure 2 shows components of an example for the system for recognition of operation noises of a vehicle comprising noise and speech feature extraction means, noise and speech recognizing means, operation noise and speech database, a recording means, vehicle component sensors and a radio transmitting device.

Figure 3 shows steps of an example of the inventive method for recognizing operation noises of a vehicle comprising detecting acoustic signals and determining whether speech signals are present as well as identification of an operation fault.

Figure 4 shows an example of the inventive method for recognizing operation noises of a vehicle comprising speech input and voice output, comprising the steps of extracting noise and speech features and running application means.

[0062] An example of the inventive system for recognition of operation noises of vehicle comprises microphones 1 installed in a vehicular cabin for detecting acoustic signals that may include speech signals and operation noise signals. The acoustic signals are transformed to electrical microphone signals and then, digitized and pre-processed by a pre-processing means 2. The pre-processing means performs a Fast Fourier Transformation and the signals coming from different microphones are synchronized by an appropriate time-delay means. Advantageously, a beamformer may be part of the pre-processing means 2.

[0063] The example also comprises a noise feature extracting means 3 and a speech feature extracting means 4. These two means are not necessarily physically separated units. By these means feature vectors are obtained corresponding to the acoustic signals detected by the microphones 1. The feature vectors comprise feature parameters that characterize the detected audio signals and are suitable for the subsequent recognition process.

[0064] Based on the feature vectors a noise and speech recognizing means 5 performs the actual recognizing process. The recognizing means makes use of a speech database 6 and an operation noise database 7. The speech database 6 comprises speech templates whereas the operation noise database 7 comprises operation noise templates. The recognizing means 5 determines the best matching template(s) for the speech signals that are present within the detected acoustic signals.

[0065] To be more specific, the templates are, according to this example, feature vectors assigned to data representations of verbal utterances. The feature vector(s) of the database that best matches the feature vector(s) obtained by analyzing the acoustic signals by the speech feature extracting means 4 is (are) determined. Thereby, the corresponding data representation is determined and the system can respond accordingly. Methods for the actual speech recognition

employing, e.g. Hidden Markov Models, are well known in the art.

[0066] Corresponding to the identified speech template a speech application means, as a telephone 8, can be run by the disclosed system. Additionally, an audio device, as a radio, can be controlled by verbal utterances of a passenger of the vehicle in this way.

[0067] If the acoustic signals detected by the microphones 1 and pre-processed by the pre-processing means 2 include operation noise signals, the associate feature vector(s) is (are) compared with the feature vectors included, as operation noise templates, in the operation noise database 7.

[0068] Depending on the determined noise template, the display device 9 shows appropriate diagnosis information. For each operation noise template or for particular classes of operation noise templates specific information can be displayed on the display device 9.

[0069] The example of the inventive system also comprises switches controlled by a control means (not shown). One switch (shown left-hand-side of the noise and speech recognition means 5 in Fig. 1) is used to input either noise feature parameters obtained by the noise feature extraction means 3 or speech feature parameters obtained by the speech feature extracting means 4 to the noise and speech recognition means 5. If, e.g., no speech signal is present, as can, e.g., be decided by the speech feature extraction means 4 or by the pre-processing means 2, only operation noise feature parameters have to be input in the recognizing means 5 that subsequently has to make use of the data input from the operation noise database 7 for the recognizing process.

[0070] Another switch allows for inputting data from the speech database 6 or the operation noise database 7 to the noise and speech recognition means 5. The switching depends on whether speech signals or operation noise signals are to be processed.

[0071] It is also possible to provide the inventive system with a push-to-talk lever that, when switched by a passenger to an "Off"-position, causes the control means to control the switches to allow connection of the recognition means 5 with the means provided for processing operation noise 3 and 7. When the push-to-talk lever is switched in an "On"-position, the control means controls the switches to allow connection of the recognition means 5 with the means provided for processing speech signals 4 and 6.

[0072] A further switch (shown on the right-hand-side of the noise and speech recognition means 5 in Fig. 1) is provided to allow running a speech application, as a telephone 8, or an application in response to operation noise recognition, as a display device 9. The switching depends either on whether the template best matching the extracted feature vector is an element of the speech database 6 or of the operation noise database 7 or on an operation of a push-to-talk lever. Different control of the above mentioned three switches as well as employment of more switching means can easily be realized by the skilled person.

[0073] As show in Fig. 2 according to another example, the system for recognition of operation noises of a vehicle comprises vehicle component sensors 10 and a recording means 11, in addition to the components shown in Fig. 1, and the application means comprise a warning means 12, a voice output 13 as well as a radio transmitting means 14.

[0074] A microphone array 1 detects acoustic signals. Whereas only one array is shown, several different ones may be installed in a vehicular cabin. The microphone array 1 comprises directional microphones pointing at different directions and converting acoustic signals into microphone signals. As in Fig. 1 the microphone signals are input in a pre-processing means 2. Both the microphone signals and the pre-processed, e.g., Fourier transformed microphone signals can be stored by a recording means 11.

[0075] Besides the microphone signals, sensor signals obtained by vehicle component sensors 10 are input in the pre-processing means 2. The sensors 10 may comprise sensors installed in the vicinity of the engine or even attached to the engine and sensors located in the individual wheel bearings. The sensor signals obtained by the vehicle component sensors 10 and the microphone signals can be synchronized by the pre-processing means 2. The sensor signals can subsequently be used by the noise and speech recognizing means 5 to improve performance and reliability of the operation noise recognizing process. If, e.g., sensor signals including information about the present engine speed are used by the recognizing means, templates of the operation noise database trained for the respective engine speed might first be compared with the presently analyzed signals, i.e., in particular, the feature vector(s) presently obtained by the feature extracting means

[0076] As in Fig. 1 a noise feature extraction means 3 analyzes the pre-processed microphones signals. The feature parameters obtained by the noise feature extraction means 3 can also be stored by the recording means 11. Thus, the recording means stores signal information at different processing stages, which is helpful in a later error analysis, e.g., during a routine inspection.

[0077] If the acoustic signals detected by the microphone array 1 contain both operation noise signals and speech signals both feature extraction means 3 and 4 may provide the recognizing means with respective feature parameters. The recognizing means determines best matching speech templates stored in the speech database 6 and in the operation noise database 7, respectively. In particular, the best matching operation noise template is preferably also stored by the recording means 11.

[0078] After operation noise signals have been processed, analyzed and recognized based on the determined best

matching operation noise template, three application means are run by the inventive system according to the present example. A warning means 12 outputs an acoustic warning, as beep sounds, if some failure in operation has been detected, i.e. if the best matching operation noise template belongs to a class of templates trained from vehicles showing some operation faults, or if the difference, in terms of some appropriate distance measure, between the extracted noise feature parameters and the feature parameters of the closest operation noise template is above a predetermined level.

[0079] Moreover, a voice output 13 is provided by which the driver can be given instructions in case of some failure. Additionally, the present example of the inventive system is equipped with a radio transmitting means 14. All data stored by the recording means 11 or input to the recording means can also be transmitted, e.g., to a service station, by the transmitting means 14.

[0080] Fig. 3 illustrates basic steps of an embodiment of the disclosed method for recognizing operation noises of a vehicle. Acoustic signals are detected 30 by microphones installed in the vehicular cabin. It is determined whether speech signals are present within the acoustic signals 31. This determination may be carried out during some signal pre-processing. In principle, speech signals are easily discriminated from noise signals by various methods known in the art.

[0081] If speech signals are present, the best matching speech template is determined 32 and subsequently, the appropriate speech application is run 34. If the acoustic signals only include noise, the best matching operation noise template is determined 33. Some of the operation noise templates represent noises of vehicles that indicate some failure, whereas other ones represent noises of faultless operation.

[0082] Depending on the identified operation noise template 35 determined to best match to the noise feature parameters obtained by analyzing noise signals either diagnosis information is displayed 36 to the driver and/or other passengers, or a warning is output 37. The latter happens, if an operation fault has been identified 35. This identification may be based on the distance of the extracted operation noise template from the best matching template. The warning can comprise acoustic warnings, as beep sounds, and visual warnings displayed on a display device.

[0083] Next, consider an example, in which both a speech input and voice output are provided as in the case of a speech dialog system. As illustrated in Fig. 4, a driver can use the speech input in demand for running audio diagnosis of operation noises of the vehicle 40. Accordingly, detected audio signals are analyzed to extract noise feature parameters 41. Subsequently, the best matching operation noise template is determined 42. If this template does not represent some operation fault 43, information about the running diagnosis can be displayed on a display device 44. If some operation fault is identified 43, the voice output prompts a warning "Operation fault" 45. The driver may advantageously be provided by further instructions as, e.g., "Stop immediately and call emergency service", in dependence on the kind of the identified operation fault.

[0084] The driver, or another passenger, may want to switch to the speech modus, after, e.g., the diagnosis has proven that operation of the vehicle is faultless. Thus, he operates a push-to-talk lever 46 to switch to the speech modus. Further utterances can demand for particular operations as dialing or controlling an entertainment system etc. Accordingly, audio signals detected after the push-to-talk lever has been switched to an "On"-position 46 are analyzed to extract speech feature parameters 47 and the best matching speech template is determined 48. Based on the identified template, i.e. data representation of the detected speech signals, some speech application is run.

[0085] All previously discussed embodiments are not intended as limitations but serve as examples illustrating features and advantages of the invention. It is to be understood that some or all of the above described features can also be combined in different ways.

Claims

1. System for automatic recognition of operation noises of a vehicle, comprising

   at least one microphone installed in a vehicular cabin for detecting acoustic signals and generating microphone signals;

   a database comprising speech templates and operation noise templates;

   feature extracting means configured to receive the generated microphone signals and to extract at least one set of noise feature parameters and/or at least one set of speech feature parameters from the generated microphone signals;

   a speech and noise recognition means configured to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters and/or to determine at least one speech template that best matches the at least one extracted set of speech feature parameters; and

   a control means configured to control the speech and noise recognition means to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters and/or

   to determine at least one speech template that best matches the at least one extracted set of speech feature parameters.

2. System according to claim 1, wherein the control means is configured to control
the feature extracting means to extract at least one set of noise feature parameters, if it controls the speech and noise recognition means to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters, and
the feature extracting means to extract at least one set of speech feature parameters, if it controls the speech and noise recognition means to determine at least one speech template that best matches the at least one extracted set of speech feature parameters.

3. System according to claim 1 or 2, wherein the control means is configured to control the speech and noise recognition means to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters, if the acoustic signals do not comprise speech signals for at least a predetermined time period.

4. System according to claim 1 or 2, further comprising a push-to-talk lever and
wherein the control means is configured to control the speech and noise recognition means to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters, if the push-to-talk lever is pushed in an "off"-position, and/or
wherein the control means is configured to control the speech and noise recognition means to determine at least one speech template that best matches the at least one extracted set of speech feature parameters, if the push-to-talk lever is pushed in an "on"-position.

5. System according to one of the preceding claims, further comprising at least one application means configured to perform applications on the basis of the at least one determined best matching speech template or the at least one determined best matching operation noise template.

6. System according to claim 5, wherein the at least one application means comprises a warning means configured to output an acoustic and/or visual and/or haptic warning, if the speech and noise recognition means is controlled to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters and if the difference between the extracted noise feature parameters and the noise feature parameters of the operation noise template exceeds a predetermined level.

7. System according to claim 5, wherein the at least one application means comprises a warning means configured to output an acoustic and/or visual and/or haptic warning, if the speech and noise recognition means is controlled to determine at least one operation noise template that best matches the at least one extracted set of noise feature parameters and if the determined operation noise template is an element of a predetermined set of particular operation noise templates indicative for operation faults.

8. System according to one of the claims 5 - 7, wherein the at least one application means comprises a wireless communication device configured to transmit data comprising the best matching operation noise template and/or the at least one extracted set of noise feature parameters and/or the generated microphone signals.

9. System according to claim 8, wherein the wireless communication device is configured to automatically transmit data comprising the best matching operation noise template and/or the at least one extracted set of noise feature parameters and/or the generated microphone signals,
if the difference between the extracted noise feature parameters and the noise feature parameters of the operation noise template determined to best match the at least one extracted set of noise feature parameters exceeds a predetermined level and/or
if the operation noise template determined to best match the at least one extracted set of noise feature parameters is an element of a predetermined set of particular operation noise templates indicative for operation faults.

10. System according to one of the claims 5 - 9, wherein the at least one application means comprise a speech output, configured to output a verbal warning,
if the difference between the extracted noise feature parameters and the noise feature parameters of the operation noise template determined to best match the at least one extracted set of noise feature parameters exceeds a predetermined level and/or
if the operation noise template determined to best match the at least one extracted set of noise feature parameters is an element of a predetermined set of particular operation noise templates indicative for operation faults.

11. System according to one of the preceding claims, further comprising at least one vehicle component sensor con-

figured to generate sensor signals; and
wherein
the speech and noise recognition means is configured to determine the at least one operation noise template that
best matches the at least one extracted set of noise feature parameters partly on the basis of the sensor signals.

12. System according to one of the preceding claims, comprising a microphone array that comprises
at least one first microphone configured for usage in common speech recognition systems and/or speech dialog
systems and/or vehicle hands-free sets and/or
at least one second microphone capable of detecting acoustic signals with frequencies below and/or above the
frequency range detected by the at least one first microphone.

13. System according to claim 12, wherein the at least one microphone array comprises at least one directional micro-
phone, in particular, more than one directional microphone pointing in different directions.

14. System according to one of the preceding claims, further comprising a beamforming means, in particular, an adaptive
beamforming means, configured to obtain beamformed microphone signals.

15. System according to one of the preceding claims, further comprising a recording means for recording the best
matching operation noise template and/or the at least one extracted set of noise feature parameters and/or the
microphone signals.

16. Method for recognizing operation noises of a vehicle comprising
providing a speech recognition system comprising a database comprising speech templates and operation noise
templates;
extracting at least one set of noise feature parameters and/or at least one set of speech feature parameters from
microphone signals generated from acoustic signals by at least one microphone installed in a vehicular cabin; and
determining at least one operation noise template that best matches the at least one extracted set of noise feature
parameters and/or determining at least one speech template that best matches the at least one extracted set of
speech feature parameters.

17. Method according to claim 16, wherein at least one set of noise feature parameters is extracted and at least one
operation noise template that best matches the at least one extracted set of noise feature parameters is determined,
if the acoustic signals do not comprise speech signals for at least a predetermined time period.

18. Method according to claim 16, wherein
at least one set of noise feature parameters is extracted and at least one operation noise template that best matches
the at least one extracted set of noise feature parameters is determined, if a push-to-talk lever is pushed in an "off"-
position and
at least one set of speech feature parameters is extracted and at least one speech template that best matches the
at least one extracted set of speech feature parameters is determined, if a push-to-talk lever is pushed in an "on"-
position.

19. Method according to one of the claims 16 - 18, wherein further
an acoustic and/or visual and/or haptic warning is output,
if the difference between the extracted noise feature parameters and the noise feature parameters of the operation
noise template determined to best match the at least one extracted set of noise feature parameters exceeds a
predetermined level or
if the operation noise template determined to best match the at least one extracted set of noise feature parameters
is an element of a predetermined set of particular operation noise templates indicative for operation faults.

20. Method according to one of the claims 16-19, wherein the best matching operation noise template and/or the at
least one extracted set of noise feature parameters and/or the generated microphone signals are transmitted by a
wireless communication device, in particular, to a service station.

21. Method according to claim 20, wherein the best matching operation noise template and/or the at least one extracted
set of noise feature parameters and/or the generated microphone signals are automatically transmitted, if the dif-
ference between the extracted noise feature parameters and the noise feature parameters of the operation noise
template determined to best match the at least one extracted set of noise feature parameters exceeds a predeter-

mined level or if the operation noise template determined to best match the at least one extracted set of noise feature parameters is an element of a predetermined set of particular operation noise templates indicative for operation faults.

22. Method according to one of the claims 16 - 21, wherein a verbal warning is output, if the difference between the extracted noise feature parameters and the noise feature parameters of the operation noise template determined to best match the at least one extracted set of noise feature parameters exceeds a predetermined level or if the operation noise template determined to best match the at least one extracted set of noise feature parameters is an element of a predetermined set of operation noise templates indicative for operation faults.

23. Method according to one of the claims 16-22, further storing the best matching operation noise template and/or the at least one extracted set of noise feature parameters and/or the microphone signals.

24. Method according to one of the claims 16 - 23, further providing at least one vehicle component sensor configured to generate sensor signals and wherein the determining of the at least one operation noise template that best matches the at least one extracted set of noise feature parameters is partly based on the sensor signals.

25. Method according to one of the claims 16-24 wherein the microphone signals are generated by at least one first microphone configured for usage in common speech recognition systems and/or speech dialog systems and/or vehicle hands-free sets and/or at least one second microphone capable of detecting acoustic signals with frequencies below and/or above the frequency range detected by the at least one first microphone.

26. Method according to one of the claims 16 - 25, wherein the microphone signals are generated by at least one directional microphone, in particular, more than one directional microphone pointing in different directions.

27. Method according to one of the claims 16-26, wherein the microphone signals are beamformed, in particular, by an adaptive beamforming means, before at least one set of noise feature parameters and/or at least one set of speech feature parameters are extracted from the microphone signals.

28. Computer program product, comprising one or more computer readable media having computer-executable instructions for performing the steps of the method according to one of the claims 16-27.

FIG. 1

FIG. 2

30
detect
acoustic
signals

32
determine
best matching
speech
template

31
speech
signals?

yes

no

33
determine
best matching
operation
noise template

run
speech
application

34

35
operation fault
identified?

no

yes

36
display
information

37
output
warning

FIG. 3

```
                    ┌─────────────────────┐
                    │   speech input:     │ ～40
                    │   "Diagnosis"       │
                    └─────────────────────┘
                              │
                              ▼
                    ┌─────────────────────┐
                    │   extract noise     │ ～41
                    │   features          │
                    └─────────────────────┘
                              │
                              ▼
                    ┌─────────────────────┐
                    │ determine best matching │ ～42
                    │ operation noise template│
                    └─────────────────────┘
                              │
  44                          ▼            43                45
   ┌─────────────┐       ◇─────────────◇        ┌──────────────────────┐
   │   display   │  no  / operation fault \ yes │    voice output      │
   │ information │◄─────│   identified?     │───►│   "Operation fault"  │
   └─────────────┘       \                 /     └──────────────────────┘
                              ◇─────◇
                              │
                              ▼
                    ┌─────────────────────┐
                    │   push-to-talk      │ ～46
                    └─────────────────────┘
                              │
                              ▼
                    ┌─────────────────────┐
                    │   extract           │ ～47
                    │   speech features   │
                    └─────────────────────┘
                              │
                              ▼
                    ┌─────────────────────┐
                    │ determine best matching │ ～48
                    │ speech template     │
                    └─────────────────────┘
                              │
                              ▼
                    ┌─────────────────────┐
                    │   run speech        │ ～49
                    │   application       │
                    └─────────────────────┘
```

FIG. 4

European Patent Office

EUROPEAN SEARCH REPORT

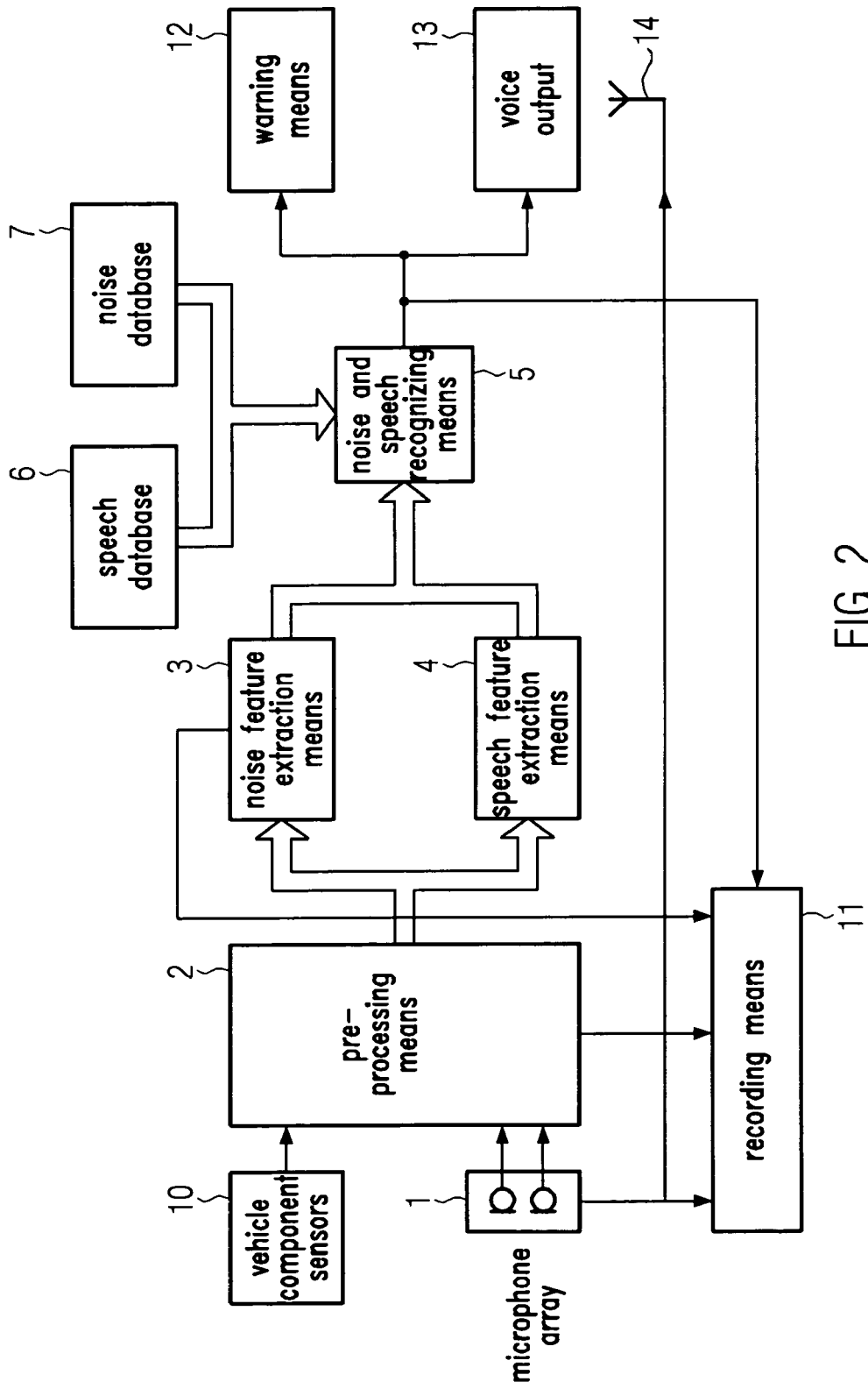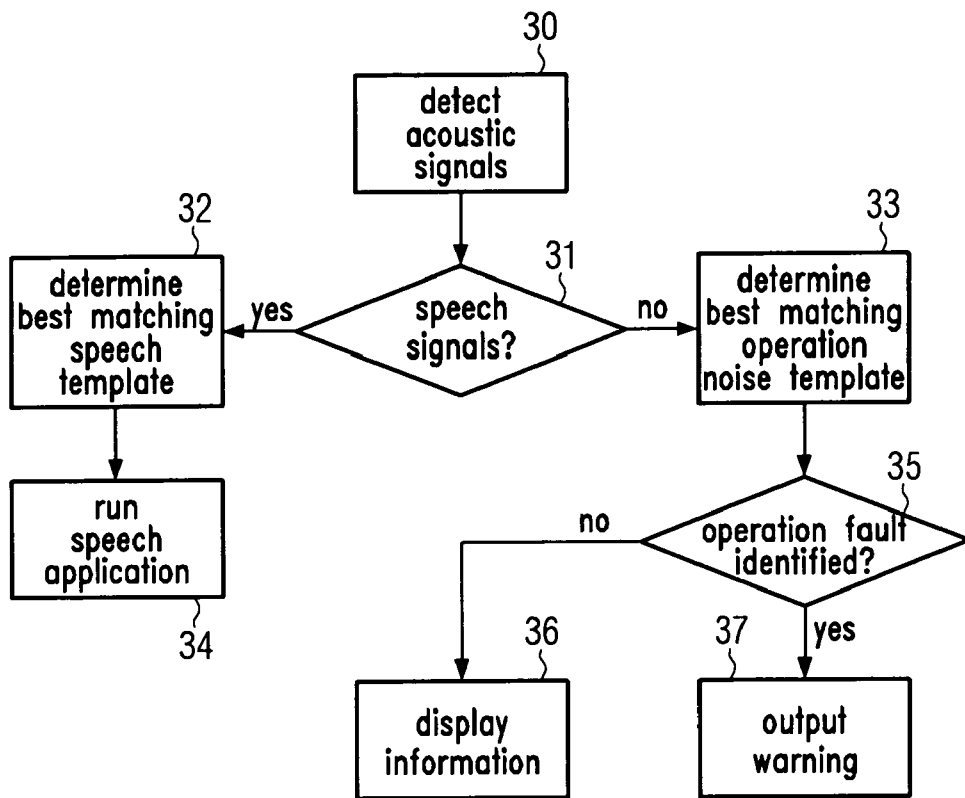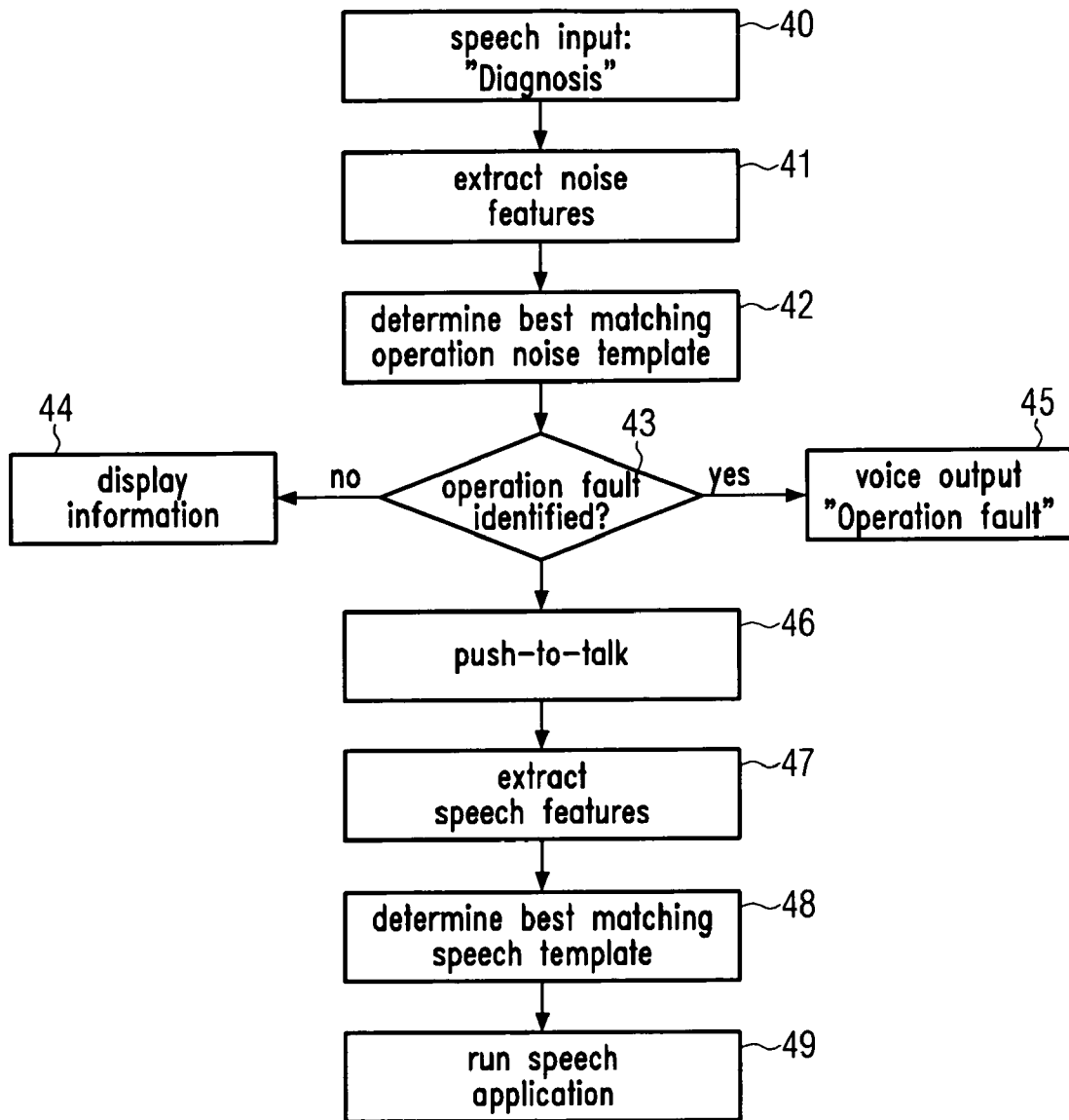## DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	DE 103 19 493 A1 (FORD GLOBAL TECHNOLOGIES, INC; FORD GLOBAL TECHNOLOGIES, LLC) 27 November 2003 (2003-11-27)	1,3-10, 14-23, 27,28	G07C5/08 G10L17/00
Y	* abstract * * column 3, paragraph 16 * * column 4, paragraph 24 * * column 4, paragraph 29 - column 5, paragraph 30 * * column 5, paragraph 32 * * column 6, paragraph 37 - column 10, paragraph 59 * * figures 1-4 *	2,11-13, 24-26	
X	DE 103 20 809 A1 (CONTI TEMIC MICROELECTRONIC GMBH) 25 November 2004 (2004-11-25)	1,5-8, 15,16, 19,20, 23,28	
	* abstract * * page 3, paragraphs 11,13,14 * * page 4, paragraphs 17,18,20,22-28 * * page 5, paragraphs 33,35,38 * * page 7, paragraphs 46,47 * * page 7, paragraph 54 - page 9, paragraph 65 * * page 9, paragraph 69 * * figures 1,5,6 *		**TECHNICAL FIELDS SEARCHED** (Int.Cl.7) G07C G10L B60R
X	US 5 748 748 A (FISCHER ET AL) 5 May 1998 (1998-05-05) * abstract * * column 5, line 65 - column 6, line 62 * * column 10, line 44 - column 12, line 4 * * figures 5-7 * -/--	1,5-7, 10,11, 16,19, 22,24,28	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
The Hague	29 July 2005	Van Der Haegen, D

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 05 00 5509

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X Y	US 2004/138882 A1 (MIYAZAWA YASUNAGA) 15 July 2004 (2004-07-15) * abstract * * page 2, paragraph 18 - paragraph 25 * * page 4, paragraph 57 - page 10, paragraph 136 * * figures 2,3,7 * -----	1,5,16, 28 2,11,24	
Y	US 2004/076305 A1 (SANTIAGO RICHARD J [US]) 22 April 2004 (2004-04-22) * abstract * * page 2, paragraph 32 - page 3, paragraph 34 * * page 3, paragraphs 37,40,41 * * page 4, paragraph 43 * * figures 4,6,7,9 * -----	12,13, 25,26	

TECHNICAL FIELDS
SEARCHED (Int.Cl.7)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
The Hague	29 July 2005	Van Der Haegen, D

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another
document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or
after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding
document

EPO FORM 1503 03.82 (P04C01)

1

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

EP 05 00 5509

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

29-07-2005

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
DE 10319493	A1	27-11-2003	US 2003216889	A1	20-11-2003
			GB 2388666	A	19-11-2003
DE 10320809	A1	25-11-2004	WO 2004101300	A1	25-11-2004
US 5748748	A	05-05-1998	DE 19531402	A1	27-02-1997
			ES 2120370	A1	16-10-1998
			FR 2738075	A1	28-02-1997
			FR 2739713	A1	11-04-1997
			GB 2305328	A	02-04-1997
			JP 3074461	B2	07-08-2000
			JP 9127957	A	16-05-1997
			JP 2000221985	A	11-08-2000
US 2004138882	A1	15-07-2004	JP 2004206063	A	22-07-2004
US 2004076305	A1	22-04-2004	AU 2003301388	A1	04-05-2004
			WO 2004036951	A2	29-04-2004

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82