

(19)



Europäisches  
Patentamt  
European  
Patent Office  
Office européen  
des brevets



(11)

**EP 2 028 651 A1**

(12)

**EUROPEAN PATENT APPLICATION**

(43) Date of publication:

**25.02.2009 Bulletin 2009/09**

(51) Int Cl.:

**G10L 21/02 (2006.01)**(21) Application number: **07114945.4**(22) Date of filing: **24.08.2007**

(84) Designated Contracting States:

**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR  
HU IE IS IT LI LT LU LV MC MT NL PL PT RO SE  
SI SK TR**

Designated Extension States:

**AL BA HR MK RS**

(71) Applicant: **Sound Intelligence B.V.**

**9712 JN Groningen (NL)**

(72) Inventors:

- **van Hengel, Peter Willem Jan**  
**9737 JP Groningen (NL)**

- **Andringa, Tjeerd Catharinus**

**9001 ZV Grou (NL)**

- **Huisman, Mark**

**9472 VP Zuidlaren (NL)**

- **Doornhein, Dimmes Abram**

**5223 MJ 's-Hertogenbosch (NL)**

- **van der Vorst, Derek**

**3743 HR Baarn (NL)**

(74) Representative: **Ketelaars, Maarten F.J.M.**

**Nederlandsch Octrooibureau**

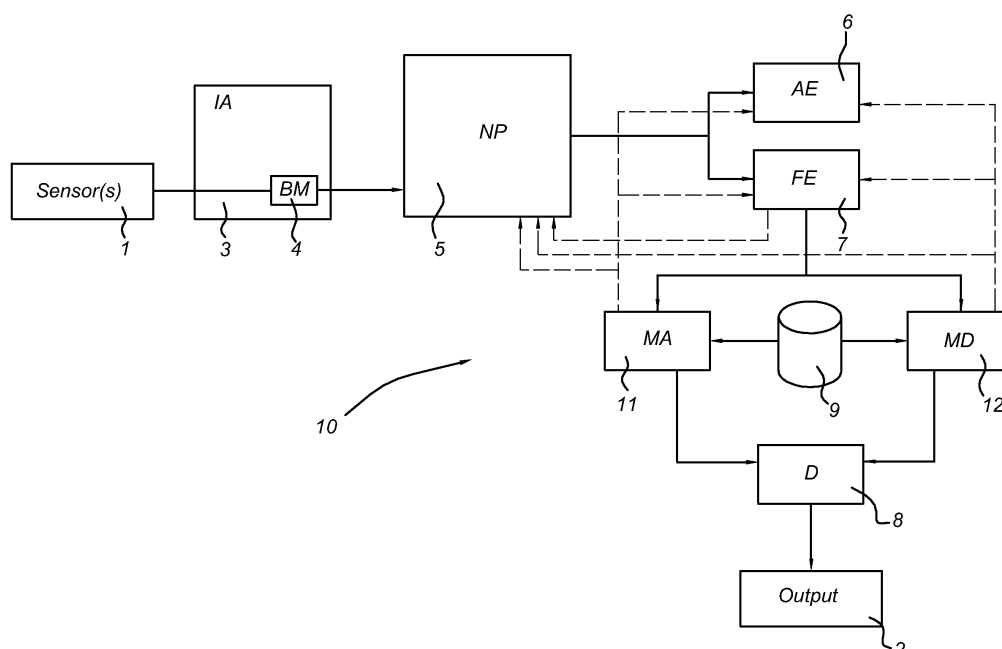
**Postbus 29720**

**2502 LS Den Haag (NL)**

(54) **Method and apparatus for detection of specific input signal contributions**

(57) Apparatus and method for detecting a single source contribution in an input signal comprising contributions from more than one source. An input analysis device (3) receives the input signal, for providing a time-frequency representation of the input signal. A neural preprocessing device (5) is connected to the input analysis device (3), for separating a foreground signal from background signals in the time-frequency representation of the input signal. A feature estimation device (7) is con-

nected to the neural preprocessing device (5) for detecting specific features in the foreground signal. A model activation device (11) is connected to the feature estimation device (7) for activating one or more of a set of models based on the detected specific features. A decision device (8) is connected to the model activation device (11) for monitoring the possible activation of a specific one of the models and generating an output based on the monitoring.

**EP 2 028 651 A1**

## Description

### Field of the invention

**[0001]** The present invention relates to a method and apparatus for detection of specific input signal contributions. More in particular, the present invention relates to an apparatus and method for detecting a single source contribution in an input signal comprising contributions from more than one source. The input signal may originate from one or more sensor devices (such as microphones), and noise may be regarded as one of the contributing sources.

### Prior art

**[0002]** European patent application EP-A-1 228 502 describes a method and apparatus for estimating frequency characteristics of an input signal. The estimation technique is based on calculating correlations of time shifted signals provided by a basilar membrane model receiving an input signal.

### Summary of the invention

**[0003]** The present invention seeks to provide an improved method and apparatus for detecting specific characteristics in signals, such as sound signals.

**[0004]** According to the present invention, an apparatus as defined above is provided, in which the apparatus comprises an input analysis device receiving the input signal, for providing a time-frequency representation of the input signal (e.g. a cochleogram), a neural preprocessing device connected to the input analysis device, for separating a foreground signal from background signals in the time-frequency representation of the input signal, a feature estimation device connected to the neural preprocessing device for detecting specific features in the foreground signal, a model activation device connected to the feature estimation device for activating one or more of a set of models (e.g. stored in a database) based on the detected specific features, and a decision device connected to the model activation device for monitoring the possible activation of a specific one of the one or more models and generating an output based on the monitoring. By separating the foreground signal from the background signals in an efficient manner, the remainder of the processing steps to detect a specific contributing source in an input signal can be made more robust, trustworthy and stable.

**[0005]** In a further embodiment, the neural preprocessing device performs the separation in foreground signal and background signals by modeling the behavior of the neural mechanisms of human hearing. This may be accomplished e.g. using modeling of the inner hair cells of the human cochlea; the outer hair cells of the human cochlea; modeling neurons in the cochlear nucleus of human brain; etc. Using modeling of mechanisms of hu-

man hearing, the eventual outcome of the detection apparatus will be more accurate and its behavior more predictable and more human-like.

**[0006]** The neural preprocessing device is, in a further embodiment, arranged to receive feedback from the feature extraction device, model activation device, and decision device, and to adapt the separation in foreground signal and background signals based on the received feedback. By using information gathered downstream the processing stages, and feeding the information back, the initial processing steps may be adapted to be more efficient in the specific situation.

**[0007]** In a further embodiment, the input analysis device comprises a basilar membrane model receiving the input signal. This basilar membrane model accurately models the behavior of the human ear, and allows to efficiently process an input signal for extracting a specific source of interest.

**[0008]** The apparatus may further comprise an auditory event device connected to the neural preprocessing device, and being arranged to determine coherent regions in the time-frequency representation of the input signal, and provide the coherent regions to the feature estimation device. This aids in detecting specific types of sources, which are likely to dominate specific regions in the time-frequency representation (e.g. harmonics).

**[0009]** The apparatus further comprises a model deactivation device connected to the feature estimation device for deactivating a model activated by the model activation device after determining that one or more specific features of the activated model are not present in the input signal. E.g. the model may have been activated because a specific feature is detected in the input signal, but when other features associated with that model are not present, it is questionable whether the model was correctly activated.

**[0010]** In a further embodiment, the model activation device is further arranged to calculate a likelihood of an activated model actually contributing to the input signal. Also, the model deactivation device may then be arranged to influence the likelihood of a specific model, and the detection device may apply rules based on the likelihood.

**[0011]** The feature estimation device may in a further embodiment be arranged to receive a feedback from the model activation device (and possibly also from the model deactivation device). Providing feedback may enhance the entire process, but may also be seen as a possibility for nesting of various devices in the apparatus. E.g. when a model is selected by a model activation device, a further feature extraction device is initiated which only searches for specific features associated with a further model closely linked to the activated model.

**[0012]** The invention may also be embodied as a method for detecting a single source contribution in an input signal comprising contributions from more than one source, comprising receiving the input signal, providing a time-frequency representation of the input signal, sep-

arating a foreground signal from background signals in the time-frequency representation of the input signal, detecting specific features in the foreground signal, activating one or more of a set of models based on the detected specific features, and monitoring the possible activation of a specific one of the one or more models and generating an output based on the monitoring. In a further embodiment, the foreground signal is separated from background signals by modeling the behavior of the neural mechanisms of human hearing. The method may further comprise receiving feedback from the detected specific features, and the activated models, and adapting the separation in foreground signal and background signals based on the received feedback. Also, the method may further comprise inputting the input signal in a basilar membrane model. In a further embodiment, the method comprises determining coherent regions in the time-frequency representation of the input signal. In an even further embodiment, the method further comprises deactivating an activated model after determining that one or more specific features of the activated model are not present in the input signal. Calculating a likelihood of an activated model actually contributing to the input signal may also be implemented in a further method embodiment. Furthermore, the method may comprise receiving feedback from the model activation for the feature extraction. These method embodiments entail similar advantages and uses as the apparatus embodiments discussed above.

**[0013]** In a further aspect, the present invention relates to a computer program product, such as a computer readable medium, comprising computer executable code, which, when loaded on a computer assembly, enables the computer assembly to execute the method according to any one of embodiments described above.

### Short description of drawings

**[0014]** The present invention will be discussed in more detail below, using a number of exemplary embodiments, with reference to the attached drawings, in which

**[0015]** Fig. 1 shows a schematic diagram of an embodiment of the apparatus of the present invention.

### Detailed description of embodiments

**[0016]** The present application describes a method and apparatus for the detection of certain (sound) sources or (sound) sources exhibiting certain characteristics in a (sound) signal that may contain (sound) signals from various sources intermixed, and possibly an unknown amount and type of background noise.

**[0017]** In Fig. 1 a schematic diagram is shown of an embodiment of the apparatus according to the present invention. One or more sensors 1 provide an input signal, comprising energy which can be described in the time and/or frequency domain. This input signal is used as input to an input analysis device (IA) 3, the function of

which will be described below. The input analysis device 3 internally uses a basilar membrane model 4 to be able to characterize the input signal using a transmission line model of the human ear, as discussed in more detail below. The output of the input analysis device 3 is a representation of the input signal in the time and frequency domain, e.g. in the form of a cochleogram.

**[0018]** The representation of the input signal is fed to a neural pre-processing device 5 (NP), the functioning of which will also be discussed in more detail below. The neural pre-processing device 5 separates a foreground signal from background signals in the representation of the input signal. This allows to more effectively detect the presence of the single source contribution in the input signal, as only the foreground signal is further inspected for certain characteristics.

**[0019]** According to embodiments of the present invention, several IA and NP devices 3, 5 can be used in series or in parallel to provide multiple representations of the input (sound) signal. This will be discussed in more detail below.

**[0020]** Using the present invention embodiments, representations of the (sound) signal can be obtained that suppress the contributions by noise and sources other than the target (sound) sources, by the use of the input analysis (IA) device 3, in combination with the neural pre-processing (NP) device 5.

**[0021]** Also these devices 3, 5 can be used to facilitate the search for cues in the (sound) signal or in a representation of the (sound) signal, that are characteristic for the target (sound) source and/or facilitate discrimination of the target (sound) source and other (sound) sources and/or noise. These devices 3, 5 can also be used to identify regions in time and/or frequency that are likely to be dominated by the target (sound) source, and/or regions that are likely to be dominated by other (sound) sources and/or noise. The likelihood that a region is dominated by the target source or by other (sound) sources and/or noise can be used by several of the other described processing and evaluation steps as a weighting function.

**[0022]** The inspection for the certain characteristics is implemented using an optional auditory event device 6 (AE), which is arranged to group regions in time and/or frequency that are likely to be dominated by the same single (sound) source (e.g. by grouping harmonics). Several possible grouping mechanisms are discussed in more detail below.

**[0023]** The inspection is implemented using a feature estimation device 7 (FE), which is arranged to compute or estimate values of characteristic features describing the (sound) signal or a part thereof. Several features that can be used to describe certain target sound sources are discussed in more detail below and the methods to accurately estimate their values in (parts) of a sound signal are described.

**[0024]** The inspection method and actual detection of the single source contribution is further implemented us-

ing a model activation device 11 (MA), model deactivation device 12 (MD) and decision device 8 (D). The decision device 8 provides an output 2, e.g. in the form of an identification of a detected single source or a trigger to a warning system.

**[0025]** The model activation device 11 is used to activate models of (sound) sources stored in a database 9. By using features extracted from the signal or a representation of the signal or parts thereof (provided by the FE device 7), models of (sound) sources are activated that exhibit those features. By comparing the values for parameters describing these features in the signal and in the model, a value for the likelihood that the model source corresponds to the (sound) source producing the input signal or contributing to the input signal is computed. By comparing the set of features describing or identifying the model source with the features that were found in the signal and/or the regions in time and frequency that contributed to the computation or estimation of the values, a reliability or quality score can be given to the model source as well.

**[0026]** The model deactivation device 12 is activated by the activation of models by the MA device 11. The MD device 12 checks the features used to describe the model associated with the source, that were not used to activate the model by the MA device 11. If such features exist, the MD device 12 checks their presence in the feature collection provided by the FE device 7. The MD device 12 may even guide the search for such features using a feedback to the FE device 7 (indicated by the dash line in Fig. 1). If features required by the model are not found in the feature collection, the MD device 12 examines the possibility that these features were obscured in the input signal by contributions to the input signal by other (sound) sources and/or noise. This is referred to as masking. If no evidence of possible masking is found, the conclusion is drawn that these features were not present in the input signal, and the MD device 12 lowers the likelihood of the associated model source accordingly.

**[0027]** The decision device 8 is used to decide which model or combination of models is actually recognized in the input. For this the decision device 8 may use the likelihood and reliability scores of the models, provided by the MA device 11 and MD device 12. The decision device 8 may use thresholds stored for these values for various models or groups of models for this task. Also this decision device 8 may use expectations based on previous input and/or previously activated models and/or previously recognized models. Also this decision device 8 may be influenced by higher level knowledge, e.g. knowledge and contextual rules stored in a context database or context model (not shown). An example of such higher level knowledge would be a language model in the case of recognition of speech, or the order of succession of alarm tones for a siren.

**[0028]** In the following, details of implementations are more broadly discussed. The IA device 3 can be used to transform the (sound) signal into a representation in

which the features used to describe the (sound) signal or parts thereof can more easily be detected and/or determined. In the case of sound signals (but not limited to these) and a detection or recognition task where human performance is the goal (but not limited to these), the logical choice is a representation of sound based on the conversion of sound into a nerve signal as it happens in the human cochlea (inner ear). A model of the (human) cochlea and of the (human) inner hair cells, can be used to come to a representation of the sound signal and/or the sound energy as a function of time and frequency, called a cochleogram. One way of modelling the cochlea is by using a so called transmission line model, which is based on equations of motion of the basilar membrane, the membrane in the cochlea on which the mechanosensitive hair cells are placed, and the surrounding fluids (see e.g. European patent application EP-A-1 228 502). The cochleogram-representation of the energy in the sound over time and frequency can contain the temporal and/or phase information on the excitation of the basilar membrane, the stimulation of the hair cells, the response of the hair cells to this stimulation, the (timed) neural spikes in the auditory nerve, the temporal information in the probability density functions for these spikes, and/or the average spike rates.

**[0029]** The NP device 5 can be used to convert the representation given by the Input Analysis Device 3 into a representation wherein certain features used to describe the target (sound) signal or parts thereof, can more easily be detected, and/or in which features not exhibited by the target (sound) signal are reduced to lower the probability of confusion, and/or in which features not exhibited by the target (sound) signal can more easily be detected in order to improve the discrimination.

**[0030]** One example of such a conversion is based on the workings of the inner hair cells in the (human) cochlea, their synapse with the auditory nerve and the cochlear nucleus nerve cells. The mechanics of these cell and synapse produce a behavior where a relatively constant background leads to a reduced response, whereas a sound that is or becomes audible above background gives an increased response. This can be modelled with a background model that is updated only when the stimulus level remains within a certain range around the previous (background) level. In general:

$$\tau \frac{dL_{bg}}{dt} + L_{bg} = L'$$

where

$$L' = S \quad \text{if} \quad S - L_{bg} \leq R$$

$$L' = L_{bg} \quad \text{if} \quad S - L_{bg} > R$$

where S is the stimulus level,  $L_{bg}$  is the background level and R is the range.

It is also possible to smooth the transition around the edges of the range R, by using:

$$L' = G(S, L_{bg})$$

where G is a smooth function which equals S for  $S - L_{bg} \leq 0$  and has  $L_{bg}$  as its limit for  $S - L_{bg} \rightarrow \infty$ .

**[0031]** It is possible to augment this procedure with an extra function which may serve to lower the background level faster than it would be based on the aforementioned formulae, which might be useful in the case of a background signal that may end more or less abruptly. The decrease in the (real) background level that this would cause can be detected and the background level compensated by using:

$$\tau_{ad} \frac{dL_{bg}}{dt} + L_{bg} = L'$$

where  $\tau_{ad}$  is shortened with respect to  $\tau$  in the original function. The shortening of  $\tau$  can be made a function of the detected fall in background level:

$$\tau_{ad} = G(S - L_{bg})$$

when

$$S - L_{bg} < -R_{ad}$$

where  $R_{ad}$  can be chosen equal to R, but it can also be give a different value.

**[0032]** By adapting the time constant  $\tau$  and/or the range R, the behaviour of the background model can be altered. Choosing a long time constant means that only slowly changing background levels are incorporated into the background model, whereas a shorter time constant means that also more rapid fluctuations may be seen as background. A good example of this would be in a system monitoring street sounds. The sound of a passing vehicle may be seen as background if the interesting sounds are screams of aggression or screams for help. If, however, the goal is to monitor traffic, the sound of the passing vehicle should not be incorporated into the background

model. Increasing the range R implies that the background model allows for larger (short term) fluctuations in level, whereas decreasing the range implies that deviations from the average level will more easily be seen as foreground sound.

**[0033]** It is possible to combine the output of background models with different time constants and/or with different ranges. If the target sound can be described in terms of its expected temporal dynamics, this procedure can be used to reduce the influence in the (neural) representation of sound sources that do not exhibit similar temporal dynamics.

**[0034]** It is also possible to simulate the way sounds grab human attention by using values for the range and time constant derived from inner hair cell and synapse models and/or measurements.

**[0035]** By comparing the energy of the foreground, at a specific location in the time-frequency plane or integrated over a certain region, with the energy of the background model an estimate can be given of the audibility or signal-to-noise ratio of certain parts of the sound. This measure can also be used to compute a reliability score for any information derived from this location or region. In the case of multiple foreground and/or background models a combination of the energies in several models can be used.

**[0036]** As an alternative and/or complementary way to achieve an improvement in the likelihood that features representative of a certain (sound) signal or parts thereof can be found, it is possible to use methods that are based on the active behaviour of outer hair cells in the (human) cochlea. These cells actively change their cell body shape in response to sound, thereby altering the (local) mechanics of the cochlea, and thereby altering the stimulus to the inner hair cells. This mechanism is under some sort of control from brain areas via an efferent nerve innervation. The effect of the so called outer hair cell motility on the cochlear mechanics can be incorporated into the model of the cochlea in several ways. It is possible to add a power source in the model at a certain location to exert a localized force. The effect of the efferent innervation can be modelled by allowing an influence by active models of certain sound sources that have been (partly) recognized or are expected by higher models (see also the detailed description on the Model Activation Device 11 and the Model Deactivation Device 12 below).

**[0037]** In the Neural Pre-processor device 5 it is possible to incorporate further neural mechanisms identified in the chain of human hearing and/or processes based on the expected action of the various neural relay stations. One possibility is to incorporate models of the different types of neurons in the Cochlear Nucleus. Three main classes of neurons have been described in the CN, defined by their behaviour and/or their response pattern:

- bushy cells, these exhibit a 'primary-like' response, that is they simply 'copy' the excitation pattern of the neurons in the auditory nerve

- stellate cells, these exhibit what is called a 'chopper' response, which can be highly tuned to a certain stimulus frequency
- octopus cells, these exhibit an 'onset' response, meaning they give a strong response to the onset of certain signals, but the response dies out even when the signal continues.

**[0038]** Incorporating these parallel neural pathways allows e.g. the Feature Extraction Device to choose the neural representation in which a feature can most reliably be determined, and/or to combine information from different neural representations. The octopus cells can be modelled by connecting the spiking outputs of various frequency channels and/or a certain frequency range, and/or by comparing, adding or multiplying the probability density functions for these spikes for various frequency channels and/or a certain frequency range and/or by comparing adding or multiplying the output of the cochlea model directly, or a measure indicating the phase of the motion, for various frequency channels and/or a certain frequency range. Other, more detailed, models for the octopus cells can also be used to generate a representation of the (sound) signal in which the onsets can more easily be found. By altering the parameters of the model, such as the frequency range or the function combining the spikes or the probability density functions, the representation of onsets can be changed.

**[0039]** The stellate cells can be modelled by connecting the spiking output of a single frequency channel or a narrow frequency range, back onto itself with a time delay. Alternatively the time signal or phase information output by the Input Analysis Device 3 can be used, or a different measure indicating the phase of the motion, and/or the probability density of the spike generation process in the synapse between the inner hair cell and the auditory nerve. By choosing a certain time delay, a certain frequency can be selected, which will give a larger response compared to other frequencies. This can be used to improve the representation of quasi-stationary frequency components in the input signal. By changing the function used to combine the output with a delayed version of itself and/or with more delayed versions of the output, the representation of these quasi-stationary components can be changed. Alternatively a measure can be made of the local instantaneous frequency of the motion based on the temporal and/or phase information in the output of the Input Analysis Device 3 or derived by another method. This measure can be used directly or in a comparison with another representation in the frequency-time plane, such as the cochleogram, to indicate the regions dominated by quasi-stationary frequency components, or to estimate the behaviour of these components with respect to e.g. frequency-development and amplitude-development over time. Fast frequency fluctuations indicative of (human) voice properties can also be determined in this manner.

**[0040]** More neural processes can be incorporated in

the Neural Pre-processing Device 5.

**[0041]** The Neural Pre-processing stage may be influenced by feedback from higher level processes, such as the feature estimation device 7, model activation device 11, model deactivation device 12, and decision device 8. Several parameters that determine the balance between the sensitivity and the reliability of the methods may be influenced by this feedback. For example the range R in the foreground background separation may initially be set to a high level to ensure that only sounds which well exceed the background level will be processed further. Once this further processing determines, however, that there may be important information missing from the foreground signal, and it could lie within the range used for the background model and therefore may have been accepted as part of the background, the range may be adapted in order to seek for this information.

**[0042]** Also higher level processing may provide expectations about the development of interesting sounds or sound properties, or of the properties of (sound) signals that are viewed as background or interference. If, for example the frequency of the target sound is expected to remain constant, the parameters of the stellate cell model may be adapted to be even more restrictive in the bandwidth of the frequency components it passes on.

**[0043]** It is possible that higher level feedback indicates that a change in parameters should be made for the analysis of the new sound coming in, or that it suggests a re-evaluation of the sound already processed with a different parameter set in order to be able to search for more and/or different cues.

**[0044]** Since this holds for all processing stages, the original sound and all intermediate representations are buffered to allow a re-examination.

**[0045]** The auditory event device 6 may be very helpful in the apparatus according to the present claim, in support of the feature extraction device 7. Auditory Events can be defined as distinct sounds as perceived by a (human) listener. In general an auditory event will be produced by a single source, although this depends on the exact definition of a (sound) source. E.g. a car can be seen as a single sound source, but it can equally well be viewed as a combination of several sound producing elements, such as wheels, doors, engine etc. In terms of the perception, an auditory event is defined as a region or combination of regions in the time-frequency plane that are grouped together to form a single internal representation. Grouping mechanisms include: common onset, common energy development over time, common fundamental frequency.

**[0046]** Firstly the Auditory Event Device 6 may determine coherent regions in the cochleogram, or in one of the alternative representations of the energy over frequency and time, that are likely to have been produced by a single sound source (or what is perceived by a naive listener as a single sound source). These regions are formed by connecting points and/or regions in the cochleogram that were defined as foreground, and share oth-

er characteristics. These characteristics may include local dominance or entrainment by the same frequency (at the same moment in time) or a continuously developing frequency (over time), a common onset and continuous energy distribution over frequency, or a common pattern of energy fluctuations over time and frequency. These coherent regions may be grouped based on a common onset, a common energy development over time, a common fundamental frequency or any other (psychophysically based) criterion.

**[0047]** The Auditory Event Device 6 may also use information from different cochleograms, produced from different sound signals, originating from different microphone signals placed within 'earshot' of one another to determine the direction from which different parts of the cochleogram energy originate. It can use the phase differences in the basilar membrane excitation, the timing differences in the neural signal and/or the level differences in any of the energy measures for this purpose. Several methods for what is called 'blind source separation' are available. The directional information can be used both in the construction of the coherent regions and in the grouping of these regions.

**[0048]** The time differences between the two sensors can be used to give an additional, spatial, dimension to the cochleogram. This can be achieved, for example, by delay-lines representing different inter-aural delays associated with different directions. These delay-lines may even incorporate level corrections to correct for the amplification, or reduction of the signal level between the two sensors. Combinations of more than two sensors can be used to improve accuracy and/or to give additional spatial dimensions.

**[0049]** Feedback from higher level processes (in particular the model activation device 11 and model deactivation device 12, see Fig. 1) may influence and/or guide the process of auditory event formation, by providing information about expected onsets, expected fundamental frequencies etc. In this way parts of regions may be removed as belonging to another sound source, parts may be added to regions as it was determined they stem from the same source, coherent regions may be removed from or added to a group based on information or expectations about sound sources, their likelihood to be present and/or their likely development.

**[0050]** The Auditory Event Device 6 may provide further processing devices (e.g. object instantiations) with a description of the auditory events found, and/or with a separate representation of these events comparable to a cochleogram or any of the other representations of the sound described so far. It may even reproduce a sound signal corresponding to the auditory event.

**[0051]** For several source models (to be described later) it may be necessary to determine the presence of certain features in the signal, in the cochleogram of the signal, in an auditory event, or any other representation of the (sound) signal. These features can be used to activate source models and/or to deactivate them. It is also

possible to derive the overall likelihood of the actual presence in the signal of the source modelled as a combination of the individual likelihoods of the features used to describe and/or define the source model.

**[0052]** The Feature Estimation Device 7 determines one or more features in the signal, in the cochleogram, in an auditory event, or in any other representation of the signal, and may even base itself on evidence found in several different representations.

**[0053]** A possible feature to be detected is the level of the signal. This feature can simply be the computed level of the parts of the energy indicated to be foreground. It is possible to integrate this measure over auditory events to give an overall level of the auditory event. It is also possible to use the reliability measures that can be determined by the Neural Pre-processing Device 5 as a weighting function in the computation of the level. Also a feedback from higher level processing (e.g. model activation device 11 or model deactivation device 12) may change the reliability scores in this computation and/or add regions to or take away regions from the computation.

**[0054]** Another possible feature to be detected is the audibility of the signal or its level compared to the background level, where background in this case refers to all other signal components present in the same time interval. This feature can simply be the ratio of the computed level of the parts of the energy indicated to be foreground and the computed level of the energy in the background model. It is possible to integrate this measure over auditory events to give an overall audibility and/or detection ability and/or reliability of information contained in the auditory event. It is also possible to use the reliability measures that can be directly determined by the Neural Preprocessing Device 5 as a weighting function in the computation of the audibility in this stage. Also a feedback from higher level processing may change the reliability scores in this computation and/or add regions to or take away regions from the computation. Another feature that can be determined is the level in any frequency band for an auditory event and/or for the energy in a foreground model and/or for the energy in a background model and/or in the original energy distribution and/or in any other representation of the (sound) signal in the time-frequency plane. A frequency band can be defined by indicating a frequency interval. Now the average energy, the total energy or any other integrated measure of the energy in the indicated frequency region can be computed. The reliability measures determined by the Neural Pre-processing Device 5 or by the Feature Estimation Device 7 itself can be used as a weighting function in this computation.

**[0055]** Another feature that can be determined is the spectral tilt of an auditory event and/or of the energy in a foreground model and/or of the energy in a background model and/or in the original energy distribution and/or in any other representation of the (sound) signal in the time-frequency plane. A high frequency region can be defined

by indicating a frequency interval, and a low frequency region can be defined in the same way. By computing the average energy, the total energy or any other integrated measure of the energy in the high frequency region and comparing this with the average energy, the total energy or any other integrated measure of the energy in the low frequency region, a spectral tilt can be computed. The reliability measures determined by the Neural Pre-processing Device 5 or by the Feature Estimation Device 7 itself can be used as a weighting function in this computation.

**[0056]** Another feature that can be determined is the presence, audibility or salience and/or the frequency of a fundamental frequency or pitch in the signal. There are several ways to determine the presence, audibility or salience and/or the frequency of a pitch in a signal. These can take any of the previous representations of the sound in the time-frequency plane and/or information about the auditory events and/or of the audibility or reliability of these events or of components of the sound defined by certain regions in the time-frequency plane or in any of the representations of the sound therein as input.

**[0057]** One possible way to compute the pitch is to look at an autocorrelation of the signal at a certain frequency or in a certain frequency band. By combining the autocorrelation information with information about the foreground background nature of the energy at any given time at the frequency or frequencies under consideration, this method can be made more robust against the influence of background signals. Information about auditory events of interest and/or auditory events possibly forming a hindrance may also improve the robustness and/or reliability. Averaging the autocorrelation over more frequencies or frequency bands or a combination of both may also improve performance. Weighting of the contributions from different frequencies or frequency regions with the average reliability of the information at that frequency, or weighting of the contributions from different frequencies at different moments in time with the reliability of the information at that point in the frequency-time plane may also improve robustness and/or reliability of the outcome. The strength of the (resulting) peak in the autocorrelation can be used to determine the strength or audibility of the pitch in the signal. The delay where the peak is found can be used to determine the frequency of the pitch estimate. Multiple peaks can be used to provide multiple pitch candidates or the pitch of multiple sources or multiple pitches in the same source.

**[0058]** Another measure for the pitch can be obtained by harmonic sieving. By computing the expected positions for harmonics for each of a dense enough base of pitch estimates and averaging or summing the levels at these positions a distribution can be created, of which the peaks form the most likely pitches present in the signal. To reduce the risk of so called octave-errors the average or sum of the levels at the harmonic positions can be corrected by the average or sum of the levels at non-harmonic positions. By choosing the non-harmonic posi-

tions to coincide with the harmonic positions of a frequency an octave lower than the frequency under study, octave errors can be reduced. Also the amount of energy explained by the pitch estimate possibly in comparison with the total energy present can be used to alter the distribution of pitch likelihoods and optimize behaviour. As with the previously described method several representations and sources of information from previously described steps can be used to improve robustness and reliability. Also feedback from higher levels may influence the parameters and/or the final distribution of likelihoods.

**[0059]** The methods can be combined to optimize the performance and/or improve the robustness and/or extend the range of reliable operation of the pitch estimation.

Several features can be based on the determination of the pitch.

**[0060]** The pitch estimate(s) can be used to derive a number of other (secondary) features, such as the distribution of energy over the harmonics (for each pitch estimate), the harmonic energy compared to the non-harmonic energy, the amount of pitches present in the signal, which can give an estimate of the number of (harmonic) sources present in the signal, the variation in frequency and/or salience of the pitch over time, the fluctuations in frequency of the individual harmonics compared to their expected frequencies derived from the pitch and the development of and/or fluctuations in the level of the individual harmonics and/or the overall harmonic pattern. As before information from different representations and/or from previously described steps and/or from feedback from higher levels or combinations of these may be used to improve the robustness and/or reliability of these features.

**[0061]** The Model Activation Device 11 can use information from the Feature Estimation Device 7 and/or from higher level processing and/or from other processing levels to activate models of (sound) signals and/or (sound) sources that may be present in the incoming (sound) signal. Such a model (stored in database 9) describes the (sound) signal or (sound) source in terms of necessary and/or possible features it may exhibit. The model database 9 contains all possible or all interesting or all expected (sound) signals or (sound) sources. A decision tree may accompany this database 9 that allows the Model Activation Device 11 to efficiently search the database 9 given a number of features that were found to be present in the incoming (sound) signal. The model may compute a likelihood that the (sound) signal or (sound) source that the model represents is present in the incoming (sound) signal based on the features found, the reliability scores of these individual features and a description of how these individual reliability scores may be combined into the overall likelihood that the (sound) signal or (sound) source that is represented by the model is present in the incoming signal. For this purpose the model may contain weighting factors for the various cues and/or a method for the computation of the overall like-



lihood.

**[0062]** The Model Activation Device 11 may actively search for features that were not originally found, by feedback to the Feature Estimation Device 7 and/or to lower level processes.

**[0063]** The models may contain features that need to be found and/or minimum reliability scores for these features that need to be exceeded in order for the model to be activated. It may also contain features which can serve to improve the overall likelihood score, but do not necessarily need to be present.

**[0064]** The Model Deactivation Device 12 searches for features that may serve to deactivate models activated either by the Model Activation Device 11 or by other means, such as feedback from higher levels and/or user control. These features may serve to differentiate models with the same or similar activating features. The features may also indicate the end of the presence of the (sound) signal or (sound) source in the incoming signal.

**[0065]** The Model Deactivation Device 12 may also deactivate models in the case of a conflict. If two or more models claim the same auditory event, or parts of it and/or the same region in the time-frequency in any of the representations of the incoming signal. Higher level feedback about the more likely of the two models and/or a decision about the accepted model may also be used to deactivate models in this case.

**[0066]** The Decision Device 8 chooses the most likely models that explain the energy and/or the features found in the incoming (sound) signal. It may not necessarily explain all the energy and/or features and/or it may chose to attribute some of the energy and/or features to a garbage model. The energy in this garbage model may be re-evaluated when more information comes in. The Decision Device 8 may contain information about the (sound) signals and/or (sound) sources that may be expected in the input and/or about likely and/or unlikely combinations of (sound) signals and/or (sound) sources. It may receive part or all of its information from user control and/or form links with knowledge systems.

**[0067]** The Decision Device 8 may pass on its interpretation of the incoming (sound) signal and/or information about certain (sound) signals and/or (sound) sources, such as start and end times and/or reliabilities of their presence or absence in the incoming (sound) signal. This interpretation and/or information may be passed on to other knowledge systems and/or learning systems and or to a database and/or to a user interface (as indicated by the output block 2 in Fig. 1).

**[0068]** The apparatus 10 in its various embodiments as described above may be used to build specific detectors. As an example of the detectors that can be made using the devices described before, an aggression detector can be constructed by using the Input Analysis Device 3 to construct a cochleogram, the Neural Preprocessing Device 5 to perform a foreground background separation based on the normal temporal development of a verbal aggression signal and/or individual aggress-

sive utterances, Feature Estimation Devices 7 to estimate the level, the audibility, the spectral tilt, the salience and frequency of the pitch, a Model Activation device 11 that activates an aggression model when these features exceed predetermined thresholds and activates a Feature Estimation Device 7 to check the level of high frequency energy and/or distortion of the spectrum and a Feature Estimation Device 7 to check the correspondence of the spectral shape with a predetermined spectral shape of aggressive utterances, and combines the levels for all features in comparison with normal values in aggressive utterances to give an overall likelihood of aggression, a Model Deactivation Device 12 that deactivates the model as soon as one or more features deviate too far from the normal values, and a Decision Device 8 that monitors the activation of the aggression model and monitors its likelihood over a predetermined time window. The Decision Device 8 may give off a warning and/or an alarm signal to another system and/or to a user interface if it has found an aggression model which was activated and retained a good likelihood for a sufficiently long time, or it may give off such a warning and/or alarm only when a certain minimum number of such 'subdetections' occur in succession within a certain time interval.

**[0069]** The aggression detector can be augmented and/or improved by using different or more representations in the Input Analysis Device 3 and/or in the Neural Preprocessing Device 5, using more features and/or a better weighted combination of these features. The advantage of this construction is that only a handful of examples of the source signal need to be available to determine the threshold and normal values for the features. This makes this setup extremely useful for constructing detectors for rare (sound) signals and/or (sound) sources. It is even possible to create a detector without any sample signals being available, if sufficient knowledge about the expected threshold and normal values of the (sound) signal and/or (sound) source to be detected exists or can be determined.

**[0070]** Further examples of detectors include, but are not limited to:

- detectors for aggression which can be used in indoor situations - jail cells and communal rooms, offices, counters, desks, interrogation rooms, emergency rooms, psychiatric wards or institutions, shelters, schools, shops, discotheques, bars, trains, busses, subways etc.

- detectors for aggression which can be used in outdoor situations - train stations, streets, car parks, parks, festivals, sporting events etc.

- detectors for panic which can be used in indoor situations - jail cells and communal rooms, offices, counters, desks, interrogation rooms, emergency rooms, psychiatric wards or institutions, shelters, schools, shops, discotheques, bars, trains, busses, subways etc.

- detectors for panic which can be used in outdoor

situations - train stations, streets, car parks, parks etc.

○ detectors for screams of pain, or other human vocalisations other than normal speech

○ detectors for the sound of vomiting, coughing, breathing, or other human sounds other than vocalisation 5

○ detectors systems for sounds indicative of vandalism, such as breaking glass, spray cans for graffiti, smashing and/or breaking sounds, impact sounds, gun shots, fireworks etc. 10

○ detectors for sounds indicative of incidents involving one or more persons, such as (street) fights, people injuring themselves or getting injured, people being otherwise in trouble 15

○ detection systems which can combine the outputs of several detectors of different types and/or from different locations

○ detection systems which can combine the outputs of several detectors from different locations and/or sensitive in different directions to determine the direction a sound or combination of sounds is coming from 20

○ detection systems which can combine the outputs of several detectors from different locations and/or sensitive in different directions to determine the location a sound or combination of sounds is coming from 25

○ tracking systems based on sound detection

○ detection systems which combine detections in a logic system which can adapt alarms signals based on a combination of alarms. This logic system can be rule based, neural network based or some other form of expert system 30

○ detectors for air plane sounds, which can be used to determine the level of air plane noise at an arbitrary location, with a far reduced influence of non-air plane sounds compared to currently available systems 35

○ detectors for vehicle noise, for example for scooters, motorbikes, heavy traffic, and/or normal cars 40

○ detection systems for multiple vehicle types and/or for traffic conditions, such as the road surface condition, the weather condition, traffic speed, traffic density

○ detection systems for sounds indicative of incidents in traffic, such as sounds associated with a collision or screeching brakes 45

○ a detector for siren sounds, either for individual siren sounds, for combinations of sirens or for sirens in general 50

○ intelligent road management systems based on sound detection, where a logic system may combine several (types of) detectors to determine traffic conditions and interface to other systems used for traffic management, such as matrix signs, traffic lights etc. 55

○ detection systems which can attribute sound levels to individual sound sources, such as an air plane noise monitoring system, a scooter noise monitoring

system, a vehicle noise monitoring system, a traffic noise monitoring system, a construction noise monitoring system, an industrial noise monitoring system which can attribute noise to different machines and/or production stages, a noise monitoring system for noise monitoring of discotheques and/or bars and/or music festivals and/or other events, or combinations of these monitoring systems

○ tracking systems based on a combination of detectors, e.g. for tracking air plane trajectories based on sound or for tracking individual vehicles based on their sound profile

○ detectors for train sounds, bus sounds, tram sounds and /or subway sounds

○ an alarm system indicating an incident or increased risk of incidents based on a.o. the sound of a passing or approaching train, bus, tram or subway

○ detectors for sounds indicative of mechanical problems with parts of a train, bus, tram or subway

○ detectors for sounds indicative of mechanical problems with mechanical systems such as lifts, escalators, doors, transport belts, gates, wheels, bearings, etc.

○ a detection system for detecting problems with wheel bearings of train wheels, mounted alongside the track, checking wheels as they pass by

○ systems based on the detection of sounds deviating from a predefined set of normal sounds for the monitoring of mechanical systems, machines, machine parts or machine complexes

○ detectors for specific abnormalities or indicators of abnormalities in signals derived from the human body other than sound, e.g. blood pressure measurements, ECG or EEG signals etc.

○ systems based on sounds and/or other signals that can be used to generate and objective diagnose of a patient's condition

○ systems based on the detection of specific sounds or other signals to monitor patients conditions

○ systems based on the detection of sounds deviating from a predefined set of normal sounds for the monitoring of patients conditions

○ detectors for sounds generated by alarm systems, e.g. a detector for fire alarm sounds, burglar alarm sounds, alarm sounds generated by medical monitoring equipment, alarm sounds generated by machine monitoring systems

○ detection systems for sounds deviating from a predefined set of normal sounds at a certain location

○ monitoring systems for the detection of non-natural sounds in a nature reserve

○ monitoring systems for sound levels in urban environments detecting non-expected sounds

○ detection systems for the detection of non-natural sounds indicative of incidents or increased risk of incidents in natural or man-made objects or constructions

○ detections systems for the detection of sounds

indicative of sliding and /or structural changes in dikes and/or embankments

- detectors for specific sounds generated by animals
- detection systems for animals or animal types based on sound detection
- detectors for specific underwater sound sources
- annotation of audio-visual content

**[0071]** Furthermore, several ways are envisaged to implement the detectors and link them to other systems:

- detection systems for sound detection implemented on special hardware, such as a DSP-platform
- detection systems for sound detection which give off a detection signal via an internet connection, which can be used e.g. by other systems to activate interface functions with other systems and/or with users
- detection systems for sound detection which give off a sound signal at detection, which can be used e.g. to alert a user or to give the microphone signal for presentation to a user or to make a recording
- detection systems which contain an audio-buffer of fixed or adjustable length which can be used to make a recording of the time interval leading up to a detection
- detection systems which give off a contact signal at detection, which can be used e.g. to activate a camera system
- detection systems which automatically log detections and/or recordings at detections and/or recordings of intervals leading up to detections
- detection systems implemented in mobile systems that can easily be installed temporarily e.g. at large events, or in other situations where, or at moments when incidents can be expected.

**[0072]** As discussed above using various exemplary embodiments, the present application describes how these methods can be used to construct a method and apparatus for the selection of the (sound) signal of a certain (sound) source or a (sound) source exhibiting certain characteristics, or the suppression of other (sound) sources and/or noise. Also the present application describes how these methods can be used to construct a method and apparatus for the recognition of (certain parts of) the (sound) signal of a certain (sound) source or a (sound) source exhibiting certain characteristics, based on reliable information about this (sound) source extracted from the total (sound) signal. Also the present application describes how several detectors and/or selectors and/or recognizers can be combined to operate on the same signal and/or on different signals to exchange and/or combine information. An example was given for the combination of two or more sound signals to derive differences in level and phase of the parts of the signal stemming from a target sound source, which can be used

for an estimation of the direction and / or position of the target sound source.

**[0073]** The present application further describes a method for reconstructing the feature set, spectrum and/or signal corresponding to a recognized or activated model. The present application describes how the use of several processing steps, to improve the reliability of information that hypotheses are based on, the careful choice of characteristic features used to describe sound sources, and the clever combination of these features, leads to detection systems that exhibit a very high degree of accuracy and reliability. The present application describes how these detection systems can be used alone or in combination to construct security and monitoring systems for a wide variety of applications.

**[0074]** Also the present application describes a method to derive model characteristics from a limited set of recordings or even from only a physical description of the signal source. The present application further describes a number of applications of the described technology. In general the technology can be used to detect and produce alarms for signals exhibiting certain specified characteristics, such as aggression, vehicles - in general or for specific types -, alarm tones, breaking glass, etc. Also the technology can be used to detect and produce alarms for signals deviating from the normal sound or sounds, where the normal sound or sounds exhibit certain specified characteristics. As mentioned before a combination of multiple detectors of both types is possible.

**[0075]** As will be apparent to the skilled person, the above method may be implemented in dedicated hardware, software or combinations of both. The software may e.g. be stored on a computer program product or computer readable medium in the form of computer executable code, which when loaded on a computer assembly, allows the computer assembly to perform the method embodiments as described above. The computer readable medium may be semiconductor based, such as a memory stick, a magnetic storage medium (e.g. a hard disk) or an optical storage medium (e.g. a CD or DVD). The computer assembly may be based on a personal computer system with associated peripheral equipment, or even a dedicated computing system, as well known to the skilled person.

## Claims

**1.** Apparatus for detecting a single source contribution in an input signal comprising contributions from more than one source, comprising

an input analysis device (3) receiving the input signal, for providing a time-frequency representation of the input signal,  
a neural preprocessing device (5) connected to the input analysis device (3), for separating a

- foreground signal from background signals in the time-frequency representation of the input signal,  
 a feature estimation device (7) connected to the neural preprocessing device (5) for detecting specific features in the foreground signal,  
 a model activation device (11) connected to the feature estimation device (7) for activating one or more of a set of models based on the detected specific features, and  
 a decision device (8) connected to the model activation device (11) for monitoring the possible activation of a specific one of the one or more models and generating an output based on the monitoring.
2. Apparatus according to claim 1, in which the neural preprocessing device (5) performs the separation in foreground signal and background signals by modeling the behavior of the neural mechanisms of human hearing.
  3. Apparatus according to claim 2, in which the neural preprocessing device (5) is arranged to receive feedback from the feature extraction device (7), model activation device (11), and decision device (8), and to adapt the separation in foreground signal and background signals based on the received feedback.
  4. Apparatus according to any one of claims 1-3, in which the input analysis device (3) comprises a basilar membrane model (4) receiving the input signal.
  5. Apparatus according to any one of claims 1-4, the apparatus further comprising an auditory event device (6) connected to the neural preprocessing device (5), and being arranged to determine coherent regions in the time-frequency representation of the input signal, and provide the coherent regions to the feature estimation device (7).
  6. Apparatus according to any one of claims 1-5, the apparatus further comprising a model deactivation device (12) connected to the feature estimation device (7) for deactivating a model activated by the model activation device (11) after determining that one or more specific features of the activated model are not present in the input signal.
  7. Apparatus according to any one of claims 1-6, in which the model activation device (11) is further arranged to calculate a likelihood of an activated model actually contributing to the input signal.
  8. Apparatus according to any one of claims 1-7, in which the feature estimation device (7) is arranged to receive a feedback from the model activation device (11).
  9. Method for detecting a single source contribution in an input signal comprising contributions from more than one source, comprising
    - receiving the input signal,  
 providing a time-frequency representation of the input signal,  
 separating a foreground signal from background signals in the time-frequency representation of the input signal,  
 detecting specific features in the foreground signal,  
 activating one or more of a set of models based on the detected specific features, and  
 monitoring the possible activation of a specific one of the one or more models and  
 generating an output based on the monitoring.
  10. Method according to claim 9, in which the foreground signal is separated from background signals by modeling the behavior of the neural mechanisms of human hearing.
  11. Method according to claim 10, further comprising
    - receiving feedback from the detected specific features, and the activated models, and  
 adapting the separation in foreground signal and background signals based on the received feedback.
  12. Method according to any one of claims 9-11, further comprising inputting the input signal in a basilar membrane model (4).
  13. Method according to any one of claims 9-12, further comprising
    - determining coherent regions in the time-frequency representation of the input signal.
  14. Method according to any one of claims 9-13, further comprising
    - deactivating an activated model after determining that one or more specific features of the activated model are not present in the input signal.
  15. Method according to any one of claims 9-14, further comprising
    - calculating a likelihood of an activated model actually contributing to the input signal.
  16. Method according to any one of claims 9-15, further comprising
    - receiving feedback from the model activation for

the feature extraction.

17. Computer readable medium comprising computer executable code, which, when loaded on a computer assembly, enables the computer assembly to execute the method according to any one of claims 9-16. 5

10

15

20

25

30

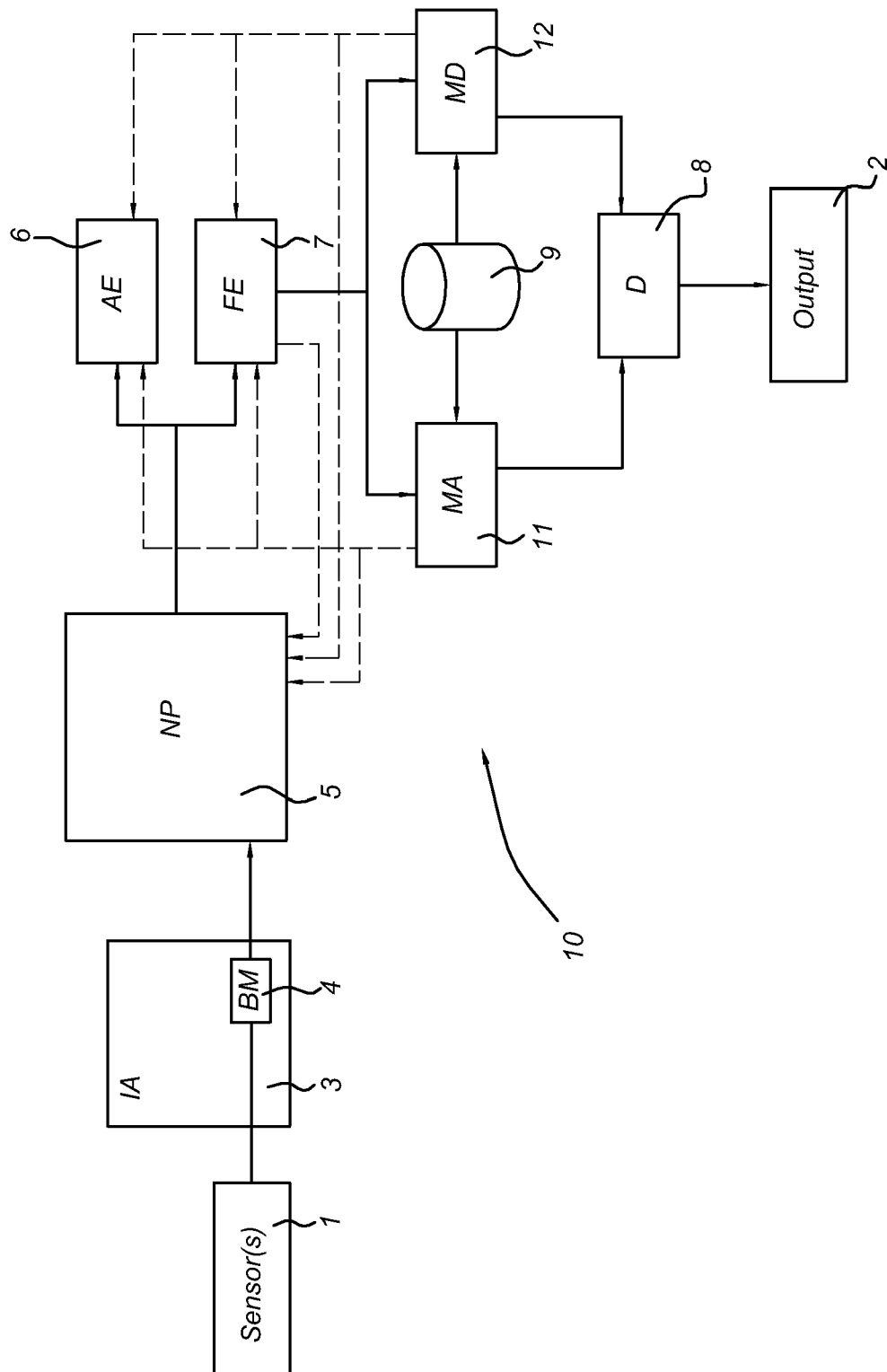
35

40

45

50

55





European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number  
EP 07 11 4945

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
D,X	WO 01/33547 A (HUQ SPEECH TECHNOLOGIES B V [NL]; ANDRINGA TJEERD CATHARINUS [NL]; DUI) 10 May 2001 (2001-05-10) * page 5, line 8 - line 11 * * page 8, line 4 - line 7 * * page 8, line 14 - line 19 * * page 15, line 24 - line 30 * * page 20, line 2 - line 11 * * page 36, line 11 - line 22 * * page 37, line 3 - line 14 * * page 45, line 14 - line 19 * * page 68, line 2 - line 10 * * page 74, line 19 - page 75, line 4 * -----	1,9,17	INV. G10L21/02
P,X	PICHEVAR ET AL: "Monophonic sound source separation with an unsupervised network of spiking neurones" NEUROCOMPUTING, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL, vol. 71, no. 1-3, 3 November 2007 (2007-11-03), pages 109-120, XP022327394 ISSN: 0925-2312 * page 110 - page 114; figure 1 * * page 118, right-hand column * -----	1,9,17	TECHNICAL FIELDS SEARCHED (IPC) G10L
X	TEDDY S D ET AL: "Model-based approach to separating instrumental music from single track recordings" CONTROL, AUTOMATION, ROBOTICS AND VISION CONFERENCE, 2004. ICARCV 2004 8TH KUNMING, CHINA 6-9 DEC. 2004, PISCATAWAY, NJ, USA, IEEE, US, 6 December 2004 (2004-12-06), pages 1808-1813Vol3, XP010815337 ISBN: 0-7803-8653-1 * page 1810 * ----- -/--	1,9,17	
The present search report has been drawn up for all claims			
Place of search Munich		Date of completion of the search 5 December 2007	Examiner RAMOS SANCHEZ, U
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ----- & : member of the same patent family, corresponding document	

1  
EPO FORM 1503 03.82 (P04C01)



European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number  
EP 07 11 4945

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	<p>DELIANG L WANG AND GUY J BROWN: "Separation of Speech from Interfering Sounds Based on Oscillatory Correlation" IEEE TRANSACTIONS ON NEURAL NETWORKS, vol. 10, no. 3, May 1999 (1999-05), pages 684-697, XP002461179 * page 685, right-hand column, paragraph 3 * * page 686, left-hand column, paragraph 1 - paragraph 3 * * page 686, right-hand column, paragraph 1 * * page 688, right-hand column, last paragraph * * page 689, right-hand column, paragraph 2 - page 690, left-hand column, paragraph 1 * * page 690, right-hand column, paragraph 2 * * page 693, right-hand column, paragraph 2 * -----</p>	1,9,17	<div>TECHNICAL FIELDS SEARCHED (IPC)</div>
The present search report has been drawn up for all claims			
Place of search <b>Munich</b>		Date of completion of the search <b>5 December 2007</b>	Examiner <b>RAMOS SANCHEZ, U</b>
<p>CATEGORY OF CITED DOCUMENTS</p> <p>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</p>		<p>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 ..... &amp; : member of the same patent family, corresponding document</p>	

1  
EPO FORM 1503 03.82 (P04C01)



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

EP 07 11 4945

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.

05-12-2007

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0133547      A	10-05-2001	AT      354849 T	15-03-2007
		AU      1740801 A	14-05-2001
		CA      2390244 A1	10-05-2001
		CN      1421030 A	28-05-2003
		DE      60033549 T2	22-11-2007
		EP      1228502 A1	07-08-2002
		JP      2003513339 T	08-04-2003
		NL      1013500 C2	08-05-2001
		US      6745155 B1	01-06-2004
-----			

**REFERENCES CITED IN THE DESCRIPTION**

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

**Patent documents cited in the description**

- EP 1228502 A [0002] [0028]