



(11) **EP 4 478 362 A1**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
18.12.2024 Bulletin 2024/51

(51) International Patent Classification (IPC):
G10L 25/03 (2013.01) **G10L 25/66** (2013.01)
G10L 25/75 (2013.01)

(21) Application number: **23461601.9**

(52) Cooperative Patent Classification (CPC):
G10L 25/03; G10L 25/66; G10L 25/75

(22) Date of filing: **12.06.2023**

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

(71) Applicant: **Vivid Mind PSA**
02-656 Warszawa (PL)

(72) Inventor: **Pluta, Adam**
02-956 Warszawa (PL)

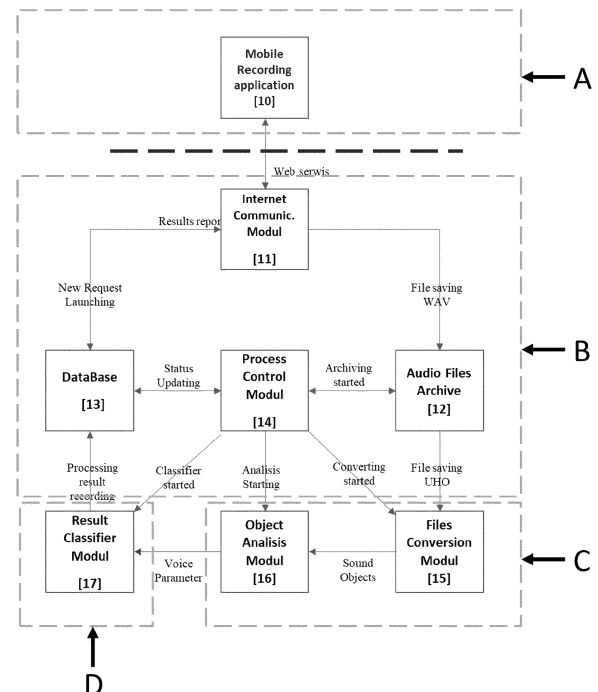
(74) Representative: **Patpol Kancelaria Patentowa Sp. z o.o.**
Nowoursynowska 162J
02-776 Warszawa (PL)

(54) **METHOD OF RECOGNITION OF THE CHARACTERISTIC FEATURES OF THE SOUND TIMBRE BASED ON SOUND OBJECTS AND SYSTEM, COMPUTER PROGRAM AND COMPUTER PROGRAM PRODUCT THEREFOR**

(57) The invention pertains to a method for recognizing the characteristics of the sound timbre based on the sound objects, in which statistical parameters are determined, after which the stages of separating sound objects into categories and determining at least one of: phase instability; the share of the energy of each category of objects in the total energy of the sound signal; harmonic frequency energy to noise ratio; amplitude slope coefficients; the slope of the graph of fundamental frequency energy and strong harmonics; local and global phase shift of each harmonic frequency; phase Fk average value, phase Fk drift; the number of sound objects and the percentage number of objects contained in the fundamental frequency and in all strong harmonic frequencies, follow.

The invention also pertains to a computer program and a computer program product comprising instructions that, when executed, perform the steps of the method for recognizing timbre characteristics based on the sound objects.

FIG 1.



EP 4 478 362 A1

Description**Field of invention**

[0001] The invention relates to a method, system, computer program and computer program product for recognizing characteristic features of sound timbre (audio signal), for example sound abnormalities, based on sound objects. The invention is applicable in the field of analysis and synthesis of acoustic signals (also referred to as sound signals or audio signals), for example, for measuring characteristic parameters of sound timbre of a human voice or characteristic sound parameters of a device, for subsequent analysis of these parameters. Parameterization of the physical phenomenon of sound emission is important in such technical fields as acoustics, mechanics, construction, aviation, astronautics, geological measurements, and many others. It also plays a role in medicine, forensics, speaker identification and speech recognition.

State of the art

[0002] In recent years, intensive development of speech analysis systems has been observed. Many programs have been created for recording speech in the form of text (the so-called "Speech-to-Text Software") or for translating text from one language to another. Researchers admit that speech and voice are also very good markers of the speaker's health. Tools are being developed to measure speech parameters and, based on a comparison with stored patterns, to diagnose various disease states.

[0003] One of the most commonly used solutions is a direct analysis of an acoustic signal. An example of such an approach is patent application US15686057A, published on March 1, 2018, "Method for evaluating a quality of voice onset of a speaker", in which the assessment of health was made by measuring the parameters of the rate of change of energy in the fundamental component and harmonic components. This study is particularly suitable for the treatment of stuttering disorders.

[0004] In patent application CN202011522343A, published on April 20, 2021, "Voice state classification method, device, electronic device and storage medium", for measuring voice parameters such as fundamental frequency jitter, amplitude shimmer, noise-to-signal ratio, a multidimensional voice processing program was applied. The advantage of the system over the patient's examination by a doctor is the objective assessment of voice parameters.

[0005] Some solutions look for voice biomarkers in spectral features and prosodic features. Melody, volume, tempo, accent, dynamics, and rhythmicity of the voice are assessed. In order to calculate these parameters, the spectrum of the acoustic signal should be determined using the Fourier Transform, or the parameters of the signal cepstrum presented in the mel scale (MFCC Mel-Frequency Cepstral Coefficients)).

[0006] An example of the use of this technology is patent application CN202110775311A, published on November 12, 2021, "A method, system and device for identifying depression based on voice analysis". This document reveals how a Digital Signal Processor with the help of an FPGA integrated circuit calculates the MFCC coefficients 1-12 and performs the classification using a decision tree algorithm.

[0007] The system presented in patent application CN202111677396A, published on April 8, 2022, "Voice analysis method and system of Parkinson's disease freezing gait symptom key characteristic parameter based on AdaBoost algorithm" works on a similar principle. In this method, the acoustic signal is analyzed using MFCC, and a CART algorithm was used to select the characteristic features. The program searches for the characteristics of Parkinson's disease in voice, which include slow speech, hoarseness, low volume, and vocal tremors.

[0008] Another solution to the problem of decomposing a signal into components, allowing to easily determine the parameters describing the acoustic signal, is presented in patent application EP16741938.1A, published on April 11, 2018, "A method and a system for decomposition of acoustic signal into sound objects, a sound object and its use". This application presents the structure of the acoustic signal spectrum, wherein the structure is illustrated using a bank of zero-phase filters with a logarithmic frequency distribution and a resolution of 48 filters per octave. From the spectrum created in this way, sound objects were created with the use of linear prediction and phase continuity tracking algorithm. The resulting sound objects very precisely reflect frequency components present in the acoustic signal, without distortion of the information contained in them about amplitude, frequency and phase.

[0009] The above-mentioned solutions known in the art offer methods for analyzing the characteristics of audio signals, allowing, among other things, to identify abnormalities in audio signals. However, none of the solutions known from the state of art use in its analysis a sound signal, recorded in the form of sound objects. Standard solutions, that do not use sound objects for analysis, require collection of large amounts of data from many samples to provide an accurate result. Analysis with the use of sound objects allows to reduce the number of samples necessary to be analyzed to obtain the appropriate parameters, allows the analysis of the sound timbre, shimmer and of the sound jitter. In addition, sound recording in the form of sound objects provides high frequency resolution of the recording, thanks to which it is possible to precisely measure the component signals of the analyzed sound. The solutions known in the art do not provide such

precise information about the sound components , in particular about the harmonic structures of the signal and their mutual relations, especially phase relations.

Summary

[0010] A method of recognizing the characteristic features of the sound timbre based on sound objects, where a sound object is analyzed, wherein a sound object is a single acoustic signal with a slowly varying amplitude and a single slowly varying frequency and continuous phase, for which sound object statistical parameters are determined:

- for each time interval, local parameters are calculated for at least one sound object: parameters of amplitude, frequency, number of measurement points;
- global parameters are calculated from local parameters,

characterized in that, after calculating the statistical parameters of sound objects, the method comprises the steps of:

- separating the sound objects into categories, wherein the categories are formed depending on frequency, wherein at least one of the categories is dependent on fundamental frequency and its associated harmonic frequencies;
- calculating categorized local and global parameters for at least one category of sound objects; and
- determining at least one of:
 - phase instability as the value of standard deviation of at least one instantaneous phase of the harmonic frequency of category of objects from at least one category from the phase of fundamental frequency,
 - the share of energy of each category of objects in the total energy of the sound signal;
 - the ratio of harmonic frequency energy to noise;
 - local and global amplitude slope coefficients;
 - directional slope coefficient of the energy graph for the fundamental frequency and strong harmonic ;
 - a local and global phase shift of each harmonic frequency relative to the fundamental frequency at the moment, when the fundamental frequency phase is 0;
 - the average value of phase Fk at the moment when the phase F1 of the fundamental frequency component has a given value; where k is the order number of the harmonic frequency;
 - the phase Fk drift, at the moment when the phase F1 of the fundamental frequency component has a given value; where k is the order number of the harmonic frequency;
 - the number of sound objects contained in the fundamental frequency and in all the strong harmonic frequencies;
 - the percentage number of objects contained in the fundamental frequency and in all the strong harmonic frequencies.

[0011] Preferably, the value of the standard deviation of at least one instantaneous phase of harmonic frequency of the category of objects is calculated by the formula

$$KHarmPhaseStandDev = \sqrt{\sum_{q=1}^Q (TabPhase[q])^2 / Q - ((\sum_{q=1}^Q TabPhase[q]) / Q)^2}$$

where TabPhase[q] is the data for the q-th measurement stored in the table, and Q is the number of measurements.

[0012] Preferably, the global parameters are calculated as a weighted average of the local parameters.

[0013] Preferably, the energy of each sound object is calculated by the formula

$$SoEnergy = \sum_{n=1}^{SoNumOfPoint} (Sp_n a)^2$$

where n is the ordinal number of the sound object point, and Sp_na is the amplitude of the sound object point for the n-th point.

[0014] Preferably, the amplitude slope coefficient is calculated by the formula

$$SumAmplean = (SumAmpPos * Num - SumAmp * SumPos) / Delta;$$

wherein

$$Num = SoNumOfPoint;$$

$$SumPos = \sum_{n=1}^{Num} Sp_{np};$$

$$SumSquarePos = \sum_{n=1}^{Num} (Sp_{np})^2;$$

$$Delta = SumSquarePos * Num - SumPos * SumPos;$$

$$SumAmp = \sum_{n=1}^{Num} Sp_{na};$$

$$SumAmpPos = \sum_{n=1}^{Num} Sp_{na} * Sp_{np}$$

where n is the ordinal number of the sound object point, SoNumOfPoint is the number of points, Sp_{na} is the amplitude of the sound object point, and Sp_{np} is the position of the sound object point.

[0015] Preferably, the average phase Fk value is calculated from the formula:

$$KHarmPhaseAvarage = (\sum_{q=1}^Q TabPhase[q])/Q;$$

where TabPhase[q] is the data for the q-th measurement stored in the table.

[0016] Preferably, the phase drift is calculated by the formula:

$$KHarmPhaseDrift = (\sum_{q=2}^Q |TabPhase[q] - TabPhase[q-1]|)/(Q-1),$$

where TabPhase[q] is the data for the q-th measurement stored in the table.

[0017] Preferably, the sound objects are divided into at least the following categories:

- fundamental frequency category;
- the category of strong harmonic frequencies, in which the share of these harmonics in the signal energy is more than 5%;
- a category of weak harmonic frequencies, in which the share of these harmonics in the signal energy is below 5%;
- category of non-harmonic signals of single frequency and amplitude;
- low frequency noise category, below 200 Hz;
- medium frequency noise category, i.e., in the range from 200 Hz to 2000 Hz;
- high-frequency noise category, i.e., in the range from 2,000 Hz to 10,000 Hz.

[0018] Preferably, after the parameter determination step, all parameters are stored and set of training input data is created for the artificial intelligence neural network.

[0019] Preferably, the time interval is 200 ms.

[0020] Preferably, the phase of the fundamental frequency component has a predetermined value.

[0021] Preferably, the phase of the fundamental frequency component is equal to 0.

[0022] A system for recognizing characteristic features of sound timbre based on sound objects, where a sound object is analyzed, wherein the sound object is a single acoustic signal with a slowly varying amplitude and a single slowly varying frequency and a continuous phase. The system is characterized by comprising:

- recording module adapted to record sounds;

- a processing module adapted to process the recorded sound so that the sound object is extracted;
- a computing module adapted to perform the above method of recognizing characteristic features of sound timbre based on the sound objects;
- a classifying module adapted to analyze the sound object based on the results obtained from the computing module.

[0023] A computer program for recognizing characteristic features of sound timbre based on sound objects, comprising instructions for performing the above method of identifying characteristic features of sound timbre based on sound objects.

[0024] A computer program product for recognizing characteristic features of sound timbre based on sound objects comprising a computer-readable code performing the steps of the above method for recognizing characteristic features of sound timbre based on sound objects.

Brief Description of the Drawings

[0025] The invention has been shown with reference to the figures on which:

Fig. 1 shows a block diagram of the analyzed person's speech processing system;

Fig. 2 is a graph of the audio signal contained in the sound object;

Fig. 3 - examples of presentation of acoustic signals using sound objects;

Fig. 4 - amplitude parameters for sound object;

Fig. 5 - frequency parameters for sound object;

Fig. 6 - phase parameters for sound object;

Fig. 7 - parameters of harmonic relations.

Description of embodiments

[0026] An exemplary application of the method according to the invention is presented below.

[0027] Fig. 1 shows a block diagram presenting the operation of the method according to the invention as well as the sequence of processes taking place in the modules of the system performing the method according to the invention.

[0028] According to the present method of recognizing the characteristic features of sound timbre, as shown in Fig. 1, using the appropriate element of the recording module (A), for example, the Mobile App (10), which can be installed on a smartphone, portable computer, notebook or other appropriate electronic device of the user, the sound is recorded in the form of a digital acoustic signal and then sent as a file without compression to the processing module (B), adapted to process the recorded sound.

[0029] The digital audio signal is stored in the .wav file in the form of a series of electrical values of the microphone signal at fixed time intervals, i.e., the so-called sampling frequency, wherein for human speech, it is sufficient to use a value of 22 050 samples per second.

[0030] In the processing module (B), the sent file goes to the Internet Network Communication Module (11), where the licensing rights of the client ordering acoustic signal processing and the data concerning the file containing information regarding what the recording is about and what file processing is to be performed are verified, all this information together can be referred to as an "order". If the client is authorized to send the file and the system is prepared to realize such processing, the authorization verification is completed correctly, the file is transferred to the audio file archive (12) and saved there.

[0031] In parallel, a new order entry is created in the database (13), for example PostgreSQL, in which order, the processing parameters are saved, including the date, time, ordering party's identifier and the type of ordered processing and, as the order is processed, the following will be added: the location of the audio file, the location of the file containing the designated sound objects (single acoustic signals with a slowly varying amplitude and a single slowly varying frequency and continuous phase), global and local parameters describing the acoustic signal, classification assessment, time of processing completion and the content of the final report sent to the client.

[0032] The audio file archive (12) and the database (13) are asynchronously connected to the process control module (14), which, after the steps of saving the file in the audio file archive (12) and creating an order record in the database (13), successively activates the computing module (C), namely sound file conversion module (15), in which at least one sound object is separated, and sound object analysis module (16) and finally the classifying module (D), also called the measurement result classifier module (17).

[0033] The sound file conversion module (15) decomposes the file, which is recorded in .wav format, from the sound file archive (12) according to the method disclosed in the application EP 3304549 A1 to create at least one sound object, and then saves this at least one sound object in the appropriate .uho format in the sound file archive (12). Information about the sampling frequency with which the signal was recorded is also sent to the .uho file. Typically, this information is in the header of the .wav file and is necessary for the correct determination of component frequencies.

[0034] The computing module (C) of Fig. 1 recognizes and is able to perform operations on various .uho recording formats, including UH00 as shown in the aforementioned patent document EP 3304549 A1.

[0035] In the next stage of the method of recognizing the characteristic features of sound timbre, the sound object analysis module (16) reads the indicated .uho file from the sound file archive (12), and then determines, for at least one sound object stored in the archive and the harmonic structures of this at least one sound object, the statistical parameters of this object, specifically local parameters: amplitude parameters, frequencies, the number of measurement points in the time interval and the length of the time interval, as well as global parameters, wherein the local parameters are calculated for time intervals and then grouped into categories according to the method of the invention, as will be described in more detail below, wherein time intervals are sections into which the sound signal recorded by the module (10) is divided. The sections may be, for example, 200 ms long (4410 samples of the acoustic signal) and local parameters are calculated for each of these sections. For example, for a 10-second recording, there are 50 time intervals, and the number of sound objects can range from 5,000 objects in the case of a low-quality recording to over 50,000 when the recording is of good quality.

[0036] The determined parameters and categories are stored in a table, which is then stored in the audio file archive (12) and, in addition, a pointer is created to the table stored in the archive (12), which is then stored in the database (13).

[0037] After module (16) conducts the analysis and creates tables stored in the archive (12), the classification module (D) reads the parameters stored in the table and compares them using artificial intelligence algorithms with data obtained from other authorized measurements, issuing a classification assessment, which is described below. The assessment will be saved in the file archive (12), after which the result of this comparison, together with the report on the execution of the order, will be sent to the user's electronic device.

[0038] Turning now to a detailed description of the method for recognizing characteristic features of sound timbre based on sound objects, the individual steps of the method are presented below.

[0039] To facilitate understanding of the equations used in the process steps, key terms of the equations that are used therein are shown below:

FileHeader -file marker;
 NumOfTrack - number of tracks;
 SampleFreq - sampling frequency;
 TrackStat - track status;
 TrackNum - track number;
 TrackNumOfSo - number of sound objects;
 TrackPos - track position;
 SoNumOfPoint - number of points in the sound object;
 SoStartPhase - initial phase in the sound object;
 SoStartPos - start position of the sound object in relation to the TrackPos position;
 SpAmp - point amplitude;
 SpTon - point tone;
 SpDeltaPos - distance from the previous point;
 k - harmonic frequency number
 n - ordinal number of the sound object point;
 m - ordinal number of the acoustic signal sample;
 Sp_n - Point number n of the analyzed sound object;
 $Sp_{n,p}$ - Position of the sound object point;
 $Sp_{n,f}$ - Frequency of the sound object point;
 $Sp_{n,\omega}$ - Pulsation of the sound object point;
 $Sp_{n,\varphi}$ - Phase of the sound object point;
 Amp_m - Amplitude of the sample m of the audio signal;
 Φ_m - Phase of the sample m of the audio signal.

[0040] In the method, after the sound is recorded, saved as a digital audio signal (as a file), the file is decomposed according to the method disclosed in application EP 3304549A1 and at least one sound object is thus obtained, further, local and global statistical parameters are determined for such a sound object.

[0041] An exemplary sound object, for example extracted by the sound file conversion module (15), is shown in Fig. 2.

[0042] Fig. 2, illustrates local parameters, determined in present method, calculated for time intervals, for each nth (n { 1... 5}) point of the analyzed sound object, in accordance with the following formulas:

$$Sp_{n,p} = Sp_{n-1,p} + Sp_{n,\Delta Pos} \quad \text{Point Position [Sample number]}$$

(continued)

$Sp_n a = Sp_n Amp$	Point Amplitude
$Sp_n f = 16.3516 * 2 ^ Sp_n Ton / (32 * 48)$	Point Frequency [Hz]
$Sp_n \omega = 2 * \pi * Sp_n f / SampleFreq$	Point Pulsation [rad/sample]
$Sp_n \varphi = Sp_{n-1} \varphi + Sp_n DeltaPoz * Sp_n \omega$	Point Phase [rad]

[0043] More specifically, Fig. 2 shows an audio signal contained in sound object, where a graph of the amplitude, frequency and phase parameters of the determined measurement points is marked point by point and a line section is used to show interpolated parameters for the signal samples between the points. The audio signal graph is defined by the following formulas:

$$Audio_m = Amp_m * \cos(\varphi_m)$$

[0044] For every n-th point $n \{ 1 \dots 4 \}$ and for each m-th sample $m \{ Sp_n p \dots Sp_{n+1} p \}$, where:

$$Amp_m = (Sp_n a * (Sp_{n+1} p - m) + Sp_{n+1} a * (m - Sp_n p)) / Sp_{n+1} DeltaPoz$$

$$\varphi_m = Sp_n a \varphi + Sp_n \omega * (m - Sp_n p)$$

[0045] The accuracy of mapping of the analyzed waveform of recorded audio signal with the above-described audio objects (calculated sample to sample) reaches 99.5%.

[0046] Described above acoustic (sound) signals, are recorded with the use of sound objects after the decomposition, as shown in Fig. 3, where three exemplary acoustic signals are illustrated:

- recording number 36 - a signal with weak vibrations, which is a sound signal coming from a person without disease symptoms;
- recording number 65 - signal with medium vibrations, which is a sound signal coming from a person suspected of having moderate dementia ailments;
- recording number 08 - a signal with strong vibrations, which is a sound signal coming from a person diagnosed with large dementia lesions.

[0047] In Fig. 3, each of the signals is presented as a graph of the audio signal over time (upper graph of each recording) and as a representation of this audio signal in the form of sound objects representing harmonic structures (the graph on the left for each recording) and non-harmonic structures and noise respectively (the graph on the right for each recording).

[0048] In an embodiment, after calculation of local parameters, it is possible to proceed to obtaining global parameters. These global parameters are calculated for the entire length of the file (for all time intervals) as weighted averages of the corresponding local parameters for all objects.

[0049] Next, after calculating the statistical parameters (local and global) of the sound objects, the sound objects are categorized according to frequency, using the parameters calculated during the previous step.

[0050] The categories into which objects are grouped are:

- the category of fundamental frequency - containing objects of fundamental frequency;
- the category of strong harmonic frequencies (low harmonics) - harmonic frequencies whose energy is greater than 5% of the energy of the entire signal;
- the category of weak harmonic frequencies (high harmonics) - containing objects with other weaker harmonic frequencies, including up to 24 frequencies;
- the category of non-harmonic signals - containing objects that are not harmonic in relation to the fundamental frequency, the energy of which exceeds 1% of the energy of the entire signal;
- the category of Low Noise (category of low-frequency noise) - containing objects with non-harmonic frequencies, the average frequency of which is lower than 200 Hz;
- the category of Medium Noise (category of medium-frequency noise) - containing objects with non-harmonic frequencies, the average frequency of which is between 200 Hz and 2 000 Hz;
- the category of High Noise (category of high-frequency noise) - containing objects with non-harmonic frequencies, in the range from 2,000 Hz to 10,000 Hz.

[0051] To reduce the number of calculations, for the first stage of the analysis one may, for example, select only objects whose frequency is less than 2,000 Hz and whose energy is not less than 1% of the energy of the entire recording, or whose length exceeds 4410 samples (200 ms).

[0052] For these several dozen sound objects, the frequencies of all pairs of objects are compared with each other (each object is successively paired with other objects), and the group of objects, that have a common divisor and the sum of their energies is the greatest, is selected. This group of objects creates a harmonic structure. The smallest divisor defines the fundamental frequency. It may happen that no object belongs to the fundamental frequency. Objects that have not been qualified to the harmonic structure will remain so-called non-harmonic objects. The objects that make up the harmonic structure will be assigned to one of the 24 harmonics. Since the distances between harmonics higher than the 24th harmonic are too small to be analyzed individually, objects with higher frequencies will be treated as noise. After separating strong objects among harmonics, weak objects will be assigned to harmonics, whose frequency coincides with the objects' harmonic frequency. Noise objects not assigned to harmonics can be divided into Low Noise, Medium Noise and High Noise.

[0053] The categorization process described above yields 24 harmonic groups, groups of non-harmonic signals with energy above 1 % of the total signal energy, and 3 types of noise based on statistical parameters present in the group of sound objects, for categories containing categorized sound objects, categorized global and local parameters are then determined. From the 24 harmonic groups, the Strong Harmonics Group is determined, which includes the fundamental frequency and those harmonics whose total energy is not less than 5% of the total energy of the recording, as well as the Weak Harmonics Group, which includes the remaining harmonics.

[0054] As in the case of statistical parameters, categorized local parameters are determined for each of the time intervals (sections) and global parameters are determined for the entire recording. The calculation of the categorized parameters for the respective categories is shown below.

[0055] In one embodiment, additionally in the method for recognizing characteristic features of the sound timbre, for the category of Low Harmonics it is possible to determine following, illustrated on the amplitude graphs of sound objects of Fig. 4, categorized local parameters according to the presented formulas:

a) Average value of the object amplitude - equal to the value of the sum of the amplitudes of all points of the sound object divided by the number of points.

$$SoAmpAverage = (\sum_{n=1}^{SoNumOfPoint} Sp_n a) / SoNumOfPoint$$

b) Energy of the object - equal to the sum of the squares of the amplitudes of all points of the objects.

$$SoEnergy = \sum_{n=1}^{SoNumOfPoint} (Sp_n a)^2$$

c) Percent value of the standard deviation of the amplitude - equal to the quotient of the square root of the difference between the object energy divided by the number of points and the square of the average amplitude value and the average object amplitude value.

$$SoAmpSDProc = SoAmpStandDev / SoAmpAvarage;$$

wherein

$$SoAmpStandDev = \sqrt{SoEnergy / SoNumOfPoint - SoAmpAvarage^2}$$

d) Percentage value of the instantaneous change in amplitude, i.e., shimmer - equal to the quotient of the quotient of the sum of the modulus of the difference in the amplitude of the adjacent points and the number of intervals and the value of the average amplitude of the object.

$$SoAmpSimmerProc = SoAmpSimmer / SoAmpAverage;$$

wherein

$$SoAmpSimmer = (\sum_{n=2}^{SoNumOfPoint} |Sp_n a - Sp_{n-1} a|) / (SoNumOfPoint - 1)$$

e) Coefficient of the slope of the amplitude representing the minimum root mean square error of the amplitude.

$$SumAmpLean = (SumAmpPos * Num - SumAmp * SumPos) / Delta;$$

wherein

$$Num = SoNumOfPoint;$$

$$SumPos = \sum_{n=1}^{Num} Sp_n p;$$

$$SumSquarePos = \sum_{n=1}^{Num} (Sp_n p)^2;$$

$$Delta = SumSquarePos * Num - SumPos * SumPos;$$

$$SumAmp = \sum_{n=1}^{Num} Sp_n a;$$

$$SumAmpPos = \sum_{n=1}^{Num} Sp_n a * Sp_n p.$$

[0056] In another embodiment, additionally in the method for recognizing characteristic features of sound timbre, for the category of Low Harmonics it is possible to determine following, illustrated on the frequency graphs of sound objects of Fig. 5, categorized global parameters according to the presented formulas:

a) The average value of the frequency of the object - the sum of the frequency values of all points of the sound object divided by the number of points.

$$SoFreqAverage = (\sum_{n=1}^{SoNumOfPoint} Sp_n f) / SoNumOfPoint$$

b) Percent value of the frequency standard deviation - equal to the quotient of the square root of the difference between the sum of the squares of the frequency of points divided by the number of points and the square of the average frequency value and the average frequency value of the object.

$$SoFreqSDProc = SoFreqStandDe / SoFreqAvarage$$

wherein

$$SoFreqSquared = \sum_{n=1}^{SoNumOfPoint} (Sp_n f)^2$$

$$SoFreqStandDe = \sqrt{SoFreqSquared / SoNumOfPoint - SoFreqAvarage^2}$$

c) The percentage value of the instantaneous frequency change, i.e., jitter - equal to the quotient of the quotient of the sum of the modules of the difference in the frequency values of the adjacent points and the number of intervals, and the value of the average frequency of the object.

$$SoFreqJitterProc = SoFreqJitter / SoFreqAvarage$$

wherein

$$SoFreqJitter = (\sum_{n=2}^{SoNumOfPoint} |Sp_n f - Sp_{n-1} f|) / (SoNumOfPoint - 1)$$

d) Coefficient of the slope of the frequency representing the minimum root mean square error of the frequency.

$$SumFreqLean = (SumFreqPos * Num - SumFreq * SumPos) / Delta$$

wherein

$$SumFreq = \sum_{n=1}^{Num} Sp_n f;$$

$$SumFreqPos = \sum_{n=1}^{Num} Sp_n f * Sp_n p .$$

[0057] The characteristic features of sound timbre recorded in the form of parameters can be used to determine the relationships between the various parameters of the signal, wherein the phase-related parameters are the preferred and most informative. Phase measurement is a very precise tool that allows, for example, to assess the quality of the speaker's voice path control. It can only be performed after properly extracting the actual harmonic frequencies present in the voice. The use of the transform for voice analysis does not give this possibility, because the harmonic frequencies of the transform do not correspond to the real harmonic frequencies. Thanks to the use of sound objects created by the Bank of Zero-Phase Filters using the phase continuity algorithm described in EP 3304549 A1, the determined harmonics correctly measure the phase of the harmonic components of the acoustic signal.

[0058] Characteristic features with respect to the parameters related to the phases are shown in Fig. 6. To determine the phase relations, it is necessary to calculate the categorized parameters analogous to those calculated above (embodiments with reference to Figs. 4 and 5) and determine the phase instability. To determine this instability, at least one instantaneous phase of the harmonic frequency from at least one category is compared to the phase of the fundamental frequency and the value of the standard deviation between these phases is determined.

[0059] If only a virtual fundamental frequency is determined in the signal, no phase measurement will be performed.

[0060] Preferably, the phase of the fundamental frequency component has a predetermined value and may be equal to 0.

[0061] In an embodiment, the measurement is performed in 200 ms sections when the phase of the fundamental frequency is 0 radians. Points in the sound object for each sound object are determined every two frequency periods of the sound object. Between two successive fundamental frequency points, there will be at least one point where the phase of the frequency will be zero. The positions of this point can be calculated from the following formula:

$$FundZeroPos = Sp_n p - Sp_n \varphi / Sp_n \omega$$

[0062] Method for calculating the phase instability for the sound object category of strong harmonic frequencies is presented below.

[0063] For each Kth harmonic frequency in the category of strong harmonic frequencies, the phase at the point zero fundamental frequency FundZeroPos can be determined, so that from all the sound objects included in this harmonic category, the sound object whose n point is closest to the designated point is selected. For this point:

$$KHarmPhaseAtPointZero = Sp_n \varphi + Sp_n \omega * (FundZeroPos - Sp_n p)$$

The fundamental frequency of the human voice is typically between 80 Hz and 250 Hz, so more than 20 points with zero value can occur in a 200 ms time interval. To determine the phase parameters in each section for each K-th harmonic frequency in the category of Strong Harmonics, the TabPhase[20] parameter was used, in which the phase values were entered in 20 consecutive points, q, zero of the fundamental frequency, where TabPhase[Q] is a table containing the determined parameters and TabPhase[20] is the data stored in the table for the twentieth parameter.

[0064] In the phase graphs of the sound objects shown in Fig. 6, phase categorized local parameters are shown,

according to the following formulas, where the phase F1 of the fundamental frequency component is assumed to be 0, but in other embodiments the phase of the fundamental frequency component may be assigned any value, which must then be taken into account when calculating other parameters:

- a) The average values of the phase of the K-th harmonic frequency when the phase of the fundamental frequency component, F1 = 0, in the category of Strong Harmonics:

$$KHarmPhaseAvarage = (\sum_{q=1}^Q TabPhase[q])/Q,$$

where TabPhase[q] is the data for the q-th measurement stored in the table, and Q is the number of measurements, where Q can be, for example, 20.

The above parameter is very important because the average phase value can be a parameter identifying the speaker.

- b) The value of the phase standard deviation of the Kth harmonic frequency in the Strong Harmonics category

$$HarmPhaseStandDev$$

$$= \sqrt{KHarmPhaseSquared/Q - KHarmPhaseAvarage^2}$$

wherein:

$$KHarmPhaseSquared = \sum_{q=1}^Q (TabPhase[q])^2;$$

- c) The value of the instantaneous phase change of the K-th harmonic frequency in the category of Strong Harmonics, when the phase of the fundamental frequency component, F1 = 0, the so-called drift:

$$KHarmPhaseDrift = (\sum_{q=2}^Q |TabPhase[q] - TabPhase[q-1]|)/(Q-1)$$

[0065] Based on the above categorized local phase parameters, the global parameters are determined; however, more information about the signal can be obtained from the locally computed parameters in 200 ms sections.

[0066] By specifying the value of the standard deviation of the phase, and thus enabling the determination of the aforementioned phase inconsistency, for any number of zero values, and thus for any number of measurements Q, the value of the standard deviation of at least one instantaneous phase of the harmonic frequency for any category of sound objects can be presented as:

$$KHarmPhaseStandDev$$

$$= \sqrt{\sum_{q=1}^Q (TabPhase[q])^2 / Q - ((\sum_{q=1}^Q TabPhase[q]) / Q)^2}$$

where TabPhase[q] is the data for the q-th measurement stored in the table, and Q is the number of measurements.

[0067] Fig. 6 shows, for three different recordings at four different times, the local harmonic components parameter's graphs, related to the fundamental frequency component. Since the graph of the acoustic signal of successive harmonics corresponds to the cosine function, the graph has a maximum value when the harmonic is in phase 0. In the subsequent parts of the chart a vertical line crossing all harmonic components marks the moment when the phase F1 of the fundamental component has a value of 0. Next to this line there are vertical lines symbolizing the zero points of the phase of successive harmonics. By measuring the distance of the vertical line and the local line sections, we can determine the phase shift of selected harmonics from the phase of the fundamental frequency. Changing phase shift of any harmonic component changes the graph of the total acoustic signal presented in the upper line of the graph. Various measured values for the phases of the harmonic components are indicated below the graphs, including phase average values, phase drifts and phase standard deviations calculated in accordance with the formulas presented in this description. A stable

value of the phase of the components guarantees the repeatability of the graph and yields information about the correct control of the speech path.

[0068] The irregular phase shift information allows the determination of the stability of phase values of the frequency components. The stable phase value guarantees the repeatability of the function, which corresponds to a sound signal graph saved in the file, and thus provides information about the correct signal path.

[0069] Determination of the non-constancy of component phases of frequency is important because changes in the relations between the phases of harmonic frequencies, especially between the phases of harmonic frequencies and the phase of the fundamental frequency, indicate whether the sound signal is controlled - a phase shift of the frequency of the fundamental component repeated in subsequent periods for the selected phase of the harmonic frequency indicates a regular and controllable sound signal, and in particular the regular and controlled operation of the sound path. A change in this shift, and more precisely in the relationship (dependence) between the phases, may indicate damage to this path or damage to the components affecting the sound generating system. Identification of the phase of a harmonic frequency that is characterized by an irregular shift can be used, for example, to identify the location of a fault.

[0070] The average value of phase shift of the selected harmonic frequency carries information about characteristic parameters of the selected fragment of the sound path and is a reference point for measuring the stability of the sound generating system. A large value of the standard deviation of phase shift of the selected harmonic frequency may be a signal that the sound path is not stable. Short-term changes (phase drift) calculated as the modulus of the phase difference at successive measurement points indicate a rapid increase or decrease in the vibration frequency of the tested harmonic and signify a deterioration of control over the analyzed fragment of the sound track.

[0071] The merging of sound objects into categories described above allows one to quickly see which categories contain how many sound objects - it is therefore possible to determine the number of sound objects contained in the fundamental frequency and in all strong harmonic frequencies, and to determine the percentage of the number of objects contained in the fundamental frequency and in all strong harmonic frequencies. Since audio objects are created when there is a step change in phase, the number of objects grouped in one category will suggest the quality and fidelity of the audio signal. The smaller the number of objects, the fewer breaks in the signal continuity occurred in the signal file, and thus the fewer breaks took place in the recorded sound.

[0072] Using methods known from the state of art, which employ signal compression or transformation, it is not possible to measure the phase parameters of the harmonic components of the sound, which carry so much information about the operation of the sound system, because the phase of components of the sound processed in this way is irretrievably distorted.

[0073] After the calculation steps and the determination of parameters, all obtained local and global parameters are placed in a table and then stored in an archive (12).

[0074] In the database (13) a pointer is created to the table stored in the archive (12), so that the database (13) serves as a register of handled orders for processed recordings.

[0075] In one embodiment, relations between the energies of the categorized sound objects are determined; for example, it is possible to analyze the energy distribution:

- between all categories - for example, the distribution can be presented in the form of a pie chart which can show, for example, the share of energy of each category of objects in the total energy of the sound signal;
- between harmonic frequencies and noise - for example, the distribution can be presented as a Harmonic to Noise Ratio (HNR);
- between strong harmonic frequencies - for example, the distribution can be presented in the form of the slope of the energy distribution line between successive harmonic frequencies;
- distribution of energy in time for selected harmonic frequencies - presented, for example, as a directional coefficient of energy slope in successive time sections.

[0076] Fig. 7 shows the distribution of energy between the harmonic and noise components of the analyzed audio signal. The pie chart on the left shows the percentage of the energy of the fundamental, low frequency strong harmonics, high frequency harmonics and other objects. The bar graph to the left of the pie chart shows the breakdown of energy for the remaining objects into low noise, medium noise, high noise and strong non-harmonic objects. Based on this graph, it is possible to determine the HNR coefficient.

[0077] The bar graph on the right in Fig. 7 graphically shows the distribution of energy between strong harmonics along with how many sound objects are contained in each harmonic. Since the sound object is interrupted when there is a step change of phase, the number of objects grouped in one harmonic indicates the quality of voice path control. The smaller the number of objects, the fewer breaks in the voice continuity. The average number of objects per harmonic is also important. As a rule, the lower the harmonic, the fewer broken objects there are. As can be seen from the examples in Fig. 7, the speech of people generating weaker vibrations is distinguished by a greater number of harmonics. In a person with severe speech disorders, the fundamental component concentrates more than 1/3 of the energy of the entire signal, and the

number of strong harmonics is small. In this case, the harmonic slope declines very strongly.

[0078] The above-mentioned information obtained by using present method of recognizing characteristic features of sound timbre on the basis of sound objects allows to assess the fidelity of the signal, for example by comparing it with other signals stored in the File Archive (12).

[0079] The local and global statistical parameters and category parameters stored in the File Archive (12) and calculated in accordance with the above-described method can be used by the classification module (D) according to the functionalities assigned to it.

[0080] The classification module (D) may be used to recognize the regularity of the signals by computer counting the objects assigned to a category or by comparing parameters of the objects to reference parameters.

[0081] Advantageously, the reference parameters can be transferred and stored in the memory of the classification module (D).

[0082] In a variant in which classification module (D) compares the parameters, the classification module (D) is informed which of the transmitted parameters are reference parameters, i.e. parameters which have been derived in accordance with the method presented herein, for:

- a sound signal known to have the correct sound characteristics, i.e. undisturbed or representing a healthy person or a properly functioning machine;
- a sound signal known to belong to an ill person or a malfunctioning machine, where it is known what the disease or damage to the machine is.

[0083] These reference parameters are provided with information, for example in the form of a flag containing an appropriate description (e.g. a healthy patient/device working properly or an ill patient/device with type failure). This stage can be defined as building an internal classifier and assigning a classification rating.

[0084] Then, any new parameters, derived from recordings of audio signals for which there is no information about the patient/device condition, can be compared in the classification module (D) with the existing reference parameters to determine whether the source of the recording (patient/device) is healthy or works fine. Thanks to this, the new parameters derived are assigned a classification rating.

[0085] In one variant, the classification module (D) may change its reference parameters, for example using artificial intelligence algorithms or neural networks, based on the provided parameters and the provided classification scores.

[0086] The method according to the invention, using sound objects for signal analysis, allows obtaining information about the signal characteristics with a relatively small number of samples, which facilitates building classifiers and further recognition of signals. In addition, the very structure of saving data from these sound objects to tables has a positive effect on the size of files recorded and saved in the database, because the connections between items in the table allow for flexible manipulation of objects and facilitate access to data related to these objects, while the size of these saved objects is small.

[0087] With the method disclosed in the present application, it is also possible to recognize the correctness of a signal, so the invention finds application in fields such as:

- 1) diagnosis of disease states affecting speech mode;
- 2) monitoring the health condition of the examined person;
- 3) voice identification of persons;
- 4) diagnosis of the state of devices by the sound emitted by the device or its component.

[0088] All embodiments of the method for recognizing timbre characteristics from sound objects are also applicable to the system, the computer program, and the computer program product.

[0089] In one embodiment, the system of the invention is implemented on a processor in any computing system such as server or PC computing system.

[0090] In another form, the invention is a computer program product comprising computer-readable instructions that cause the processor to perform the method of recognizing the characteristic features of sound timbre according to the invention.

[0091] Each of the schematic blocks of the system for recognizing characteristic features of sound timbre based on sound objects, as well as each of the steps of the method for recognizing characteristic features of sound timbre based on sound objects, may be implemented by computer program instructions. These instructions may be provided to a general purpose computer processor, a special purpose computer or other programmable data processing device such that the instructions that are executed by the computer processor or other programmable data processing device enable

implementation of the functions defined in the system diagram and method steps.

[0092] Aspects of the present invention may be implemented by a computer or devices such as a CPU (central processing unit), MPU (microprocessor unit) or MCU (microcontroller unit) that read and execute a computer program product stored in the device memory to perform the functions of the embodiments described above. Aspects of the present invention may also be implemented by a method the steps of which are performed by a computer of the system or device by, for example, reading and executing a program stored on a memory device to perform the functions of the embodiments described above. In this regard, the computer program product is provided to the computer, for example, over a network or other type of recording medium serving as a storage device. The computer program product of the invention also includes a durable machine-readable medium.

[0093] Embodiments are provided herein as non-limiting indications of the invention only and are not intended to limit in any way the scope of protection that is defined by the claims, and terms such as sound timbre refer to sound characteristics and may also be referred to as components, intensity or harmonics, without affecting the scope of the invention. It should be understood that each technical solution used in the invention can be implemented using equivalent technologies without going beyond the scope of protection.

Claims

1. A method of recognizing the characteristic features of sound timbre based on sound object, where a sound object is analyzed, wherein the sound object is a single acoustic signal with a slowly varying amplitude and a single slowly varying frequency and continuous phase, for which sound object statistical parameters are determined:

- for each time interval, local parameters are calculated for at least one sound object:
parameter of amplitude, frequency, number of measurement points;
- global parameters are calculated from local parameters;

characterized in that after calculating the statistical parameters of the sound objects, the method comprises the steps of:

- separating the sound objects into categories, wherein the categories are formed depending on the frequency, wherein at least one of the categories is dependent on the fundamental frequency and its associated harmonic frequencies;
- calculating categorized local and global parameters for at least one category of sound objects; and
- determining at least one of:

- o phase instability as the value of the standard deviation of at least one instantaneous phase of the harmonic frequency of category of objects of at least one category from fundamental frequency phase;
- o the share of the energy of each category of objects in the total energy of the sound signal;
- o the ratio of harmonic frequency energy to noise;
- o local and global amplitude slope coefficients;
- o directional slope coefficient of the energy graph of the fundamental frequency and strong harmonic;
- o a local and global phase shift of each harmonic frequency relative to the fundamental frequency at the moment when the fundamental frequency phase is 0;
- o the average value of phase F_k at the moment when the phase F₁ of the fundamental frequency component has a given value; where k is the order number of the harmonic frequency
- o the phase F_k drift at the moment when the phase F₁ of the fundamental frequency component has a given value; where k is the order number of the harmonic frequency;
- o the number of sound objects contained in the fundamental frequency and in all strong harmonic frequencies;
- o the percentage number of objects contained in the fundamental frequency and in all strong harmonic frequencies.

2. The method according to claim 1, **characterized in that** the value of the standard deviation of at least one instantaneous phase of the harmonic frequency of category of objects is calculated by the formula

$$KHarmPhaseStandDev = \sqrt{\sum_{q=1}^Q (TabPhase[q])^2 / Q - ((\sum_{q=1}^Q TabPhase[q]) / Q)^2}$$

where TabPhase[q] is the data for the q-th measurement stored in the table, and Q is the number of measurements.

3. The method according to any one of claim 1 or claim 2, **characterized in that** the global parameters are calculated as a weighted average of the local parameters.
4. The method according to claim 1, **characterized in that** the energy of each sound object is calculated by a formula:

$$SoEnergy = \sum_{n=1}^{SoNumOfPoint} (Sp_n a)^2$$

where n is the ordinal number of the sound object point, and Sp_na is the amplitude of the sound object point for the n-th point.

5. The method according to claim 1, **characterized in that** the amplitude slope coefficient is calculated by the formula

$$SumAmpLean = (SumAmpPos * Num - SumAmp * SumPos) / Delta;$$

wherein

$$Num = SoNumOfPoint;$$

$$SumPos = \sum_{n=1}^{Num} Sp_n p;$$

$$SumSquarePos = \sum_{n=1}^{Num} (Sp_n p)^2;$$

$$Delta = SumSquarePos * Num - SumPos * SumPos;$$

$$SumAmp = \sum_{n=1}^{Num} Sp_n a;$$

$$SumAmpPos = \sum_{n=1}^{Num} Sp_n a * Sp_n p$$

where n is the ordinal number of the sound object point, SoNumOfPoint is the number of points, Sp_na is the amplitude of the sound object point, and Sp_np is the position of the sound object point.

6. The method according to claim 1, **characterized in that** the average phase Fk value is calculated from the formula:

$$KHarmPhaseAvarage = (\sum_{q=1}^Q TabPhase[q]) / Q;$$

where TabPhase[q] is the data for the q-th measurement stored in the table.

7. The method according to claim 1, **characterized in that** the phase drift is calculated by the formula:

$$KHarmPhaseDrift = (\sum_{q=2}^Q |TabPhase[q] - TabPhase[q-1]|)/(Q-1),$$

where TabPhase[q] is the data for the q-th measurement stored in the table.

8. The method according to any one of claims 1-7, **characterized in that** the sound objects are divided into at least the following categories:

- fundamental frequency category;
- the category of strong harmonic frequencies, in which the share of these harmonics in the signal energy is more than 5%;
- a category of weak harmonic frequencies, in which the share of these harmonics in the signal energy is below 5%;
- category of non-harmonic signals of single frequency and amplitude,
- low frequency noise category, below 200 Hz;
- medium frequency noise category, i.e. in the range from 200 Hz to 2000 Hz;
- high-frequency noise category, i.e. in the range from 2,000 Hz to 10,000 Hz.

9. The method according to any one of claims 1-8, **characterized in that** after the determination step, all parameters are stored and a set of training input data is created for the artificial intelligence neural network.

10. The method according to any one of claims 1-9, **characterized in that** the time interval is 200 ms.

11. The method according to any one of claims 1-10, **characterized in that** the phase of the fundamental frequency component has a predetermined value.

12. The method according to claims 11, **characterized in that the** phase of the fundamental frequency component is equal to 0.

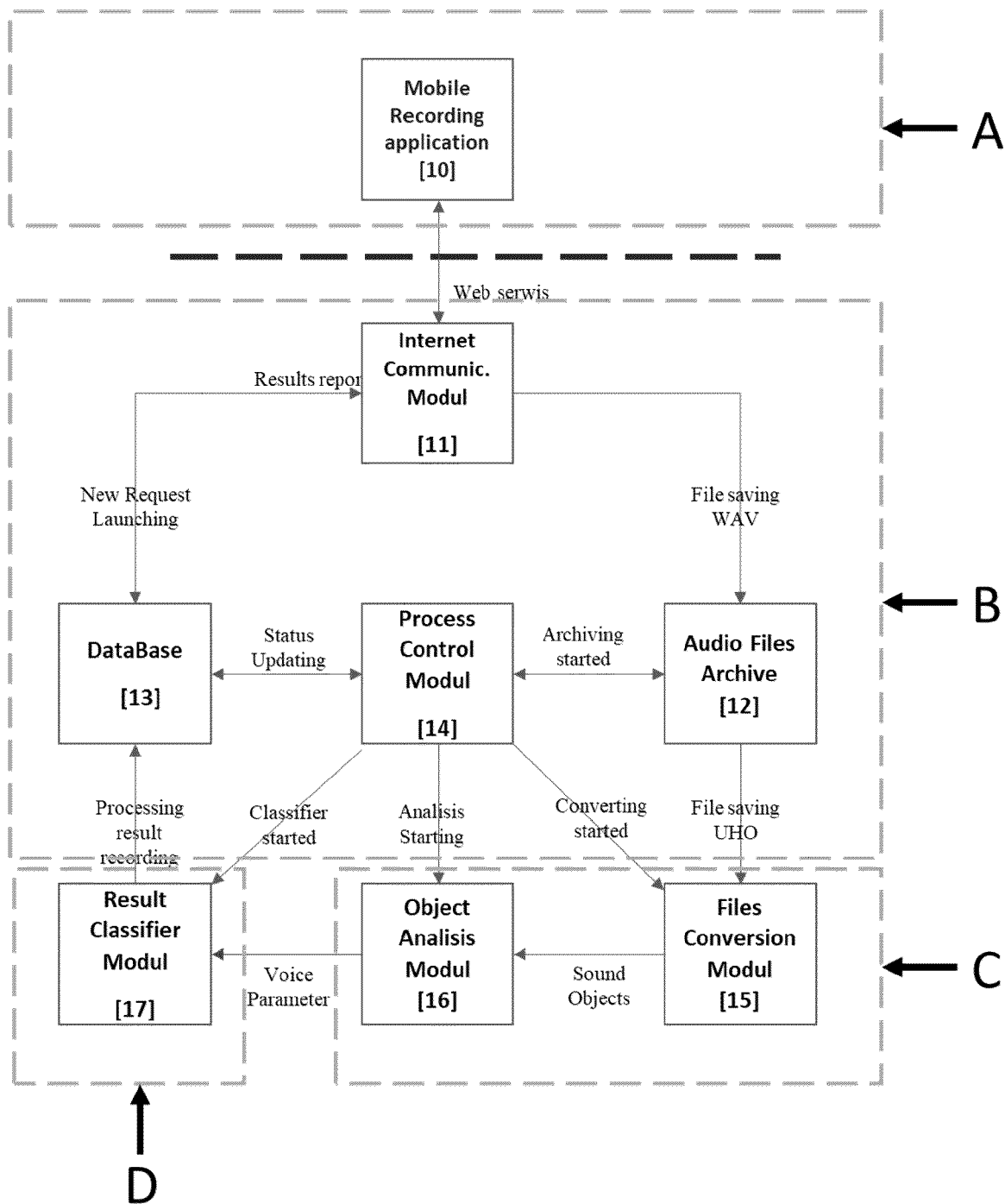
13. A system for recognizing characteristic features of sound timbre based on sound objects, where a sound object is analyzed, wherein the sound object is a single acoustic signal with a slowly varying amplitude and a single slowly varying frequency and continuous phase, **characterized by** comprising:

- recording module adapted to record sounds;
- a processing module adapted to process the recorded sound so that the sound object is extracted;
- a computing module adapted to perform the method of claims 1 to 12;
- a classifying module adapted to analyze the sound object based on the results obtained from the calculation module.

14. A computer program for recognizing characteristic features of sound timbre based on sound objects, comprising instructions for performing the method of any one of claims 1 to 12.

15. A computer program product for recognizing characteristic features of sound timbre based on sound objects comprising a computer-readable code performing the steps of the method of any of claims 1 to 12.

FIG 1.



Sound Object example

Point Number	Position [Sample]	Amplitude	Frequency [Filter No * 32]	Phase [Degrees * 100]
1	13317	28	6916	4078
2	13437	28	6895	4013
3	13557	53	6873	3237
4	13677	63	6875	2526
5	13797	36	6889	2266

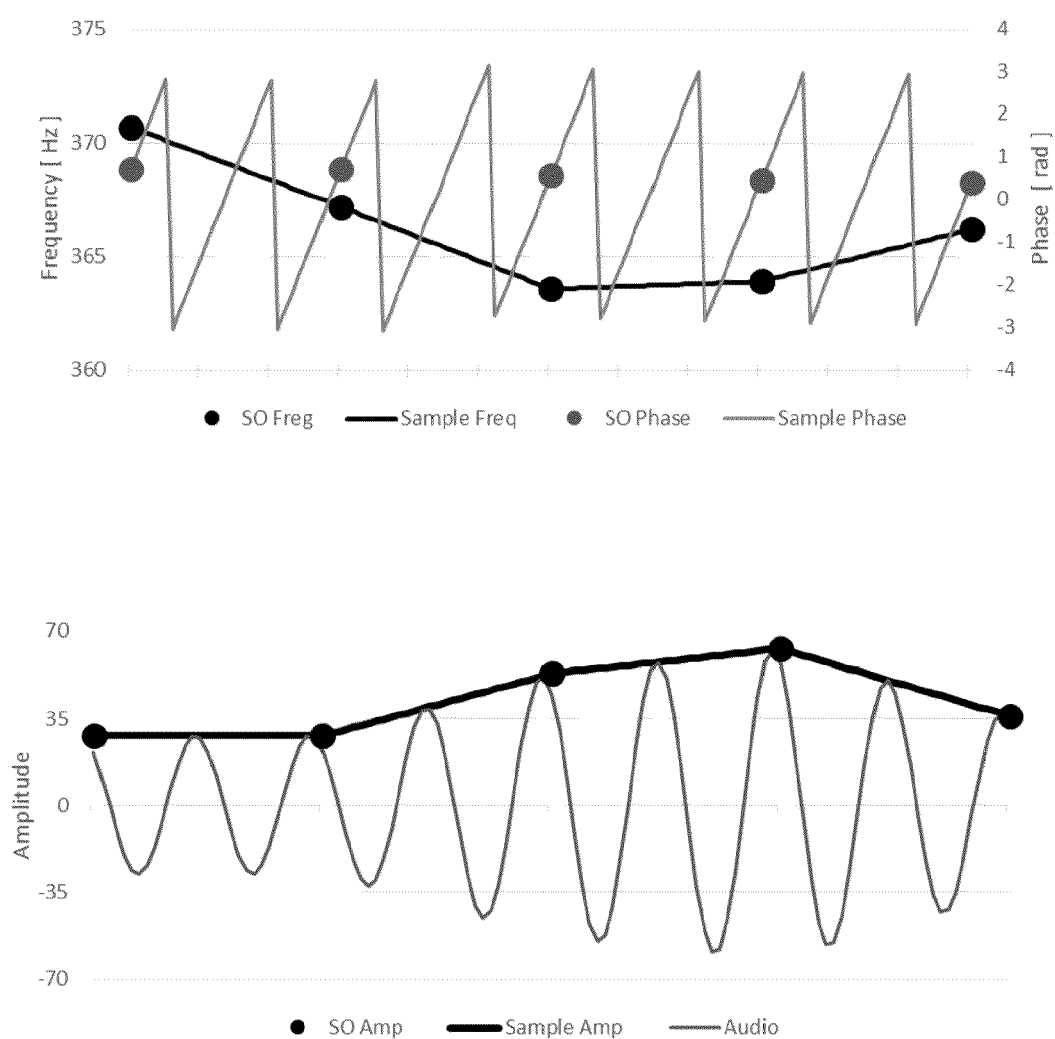
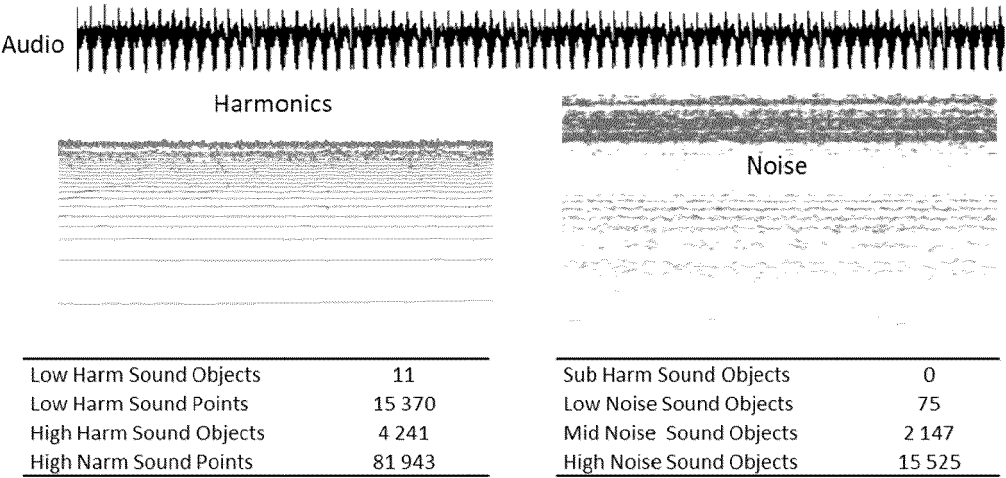
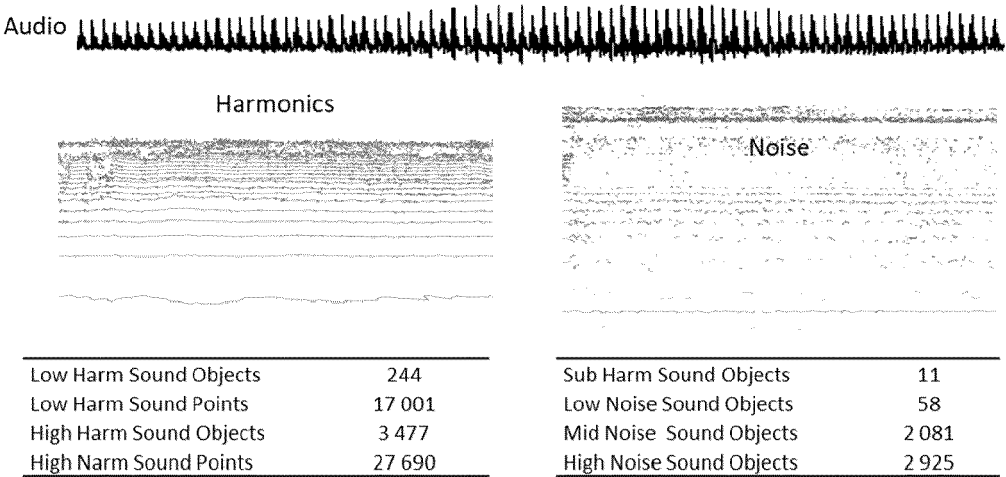


Fig. 2

Little vibrations (36)



Medium vibrations (65)



High vibrations (8)

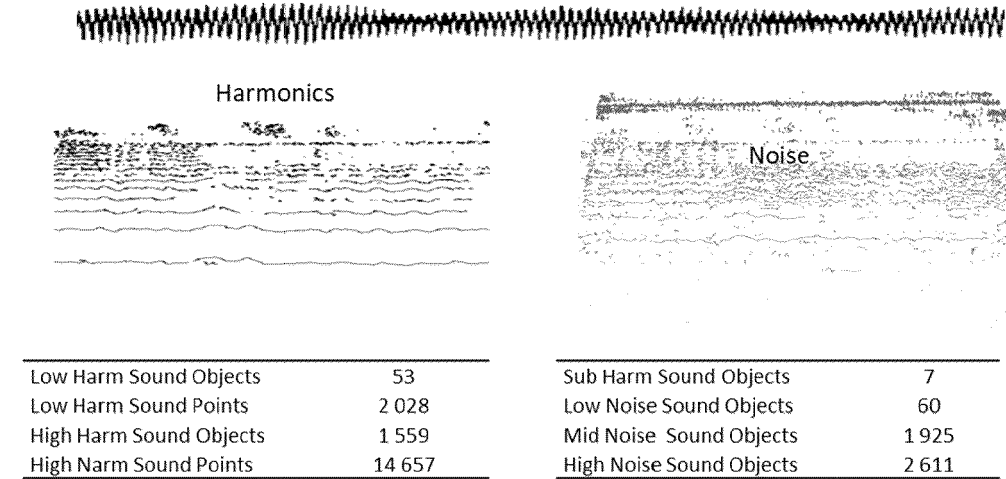
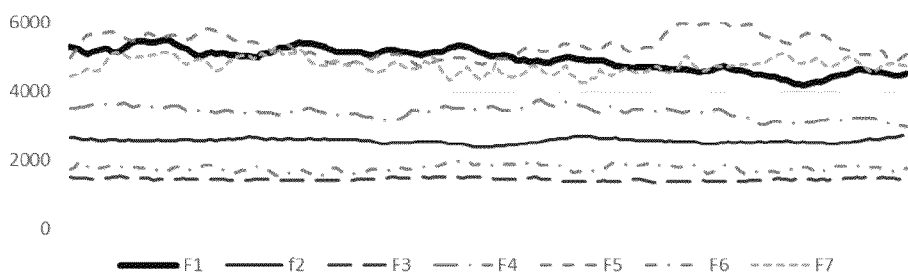


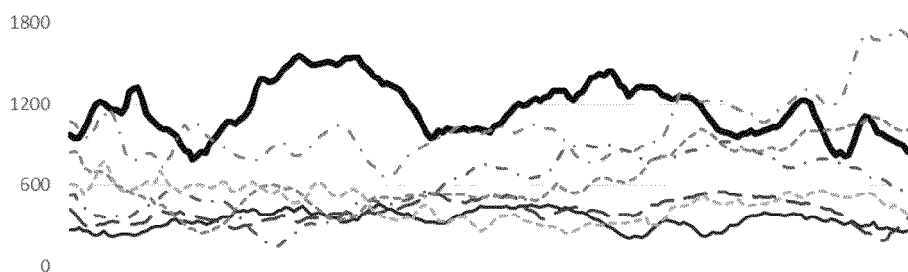
Fig 3.

Little vibrations



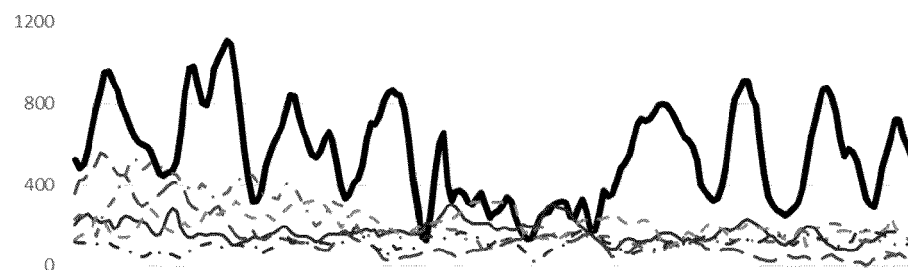
	F1	F2	F3	F4	F5	F6	F7
AmpAv	4904	2473	1342	3275	5382	2144	5028
AmpSD	6,1%	5,5%	6,2%	8,7%	10,9%	11,1%	6,8%
Shimmer	0,46%	0,35%	0,40%	0,42%	0,45%	0,69%	0,52%
Lean	-0,55%	-0,14%	-0,37%	0,39%	0,90%	2,16%	0,66%

Medium vibrations



	F1	F2	F3	F4	F5	F6	F7
AmpAv	979	291	324	411	464	899	601
AmpSD	24,0%	24,1%	23,8%	36,7%	33,2%	28,5%	23,4%
Shimmer	1,63%	2,23%	1,94%	1,85%	2,00%	1,53%	1,93%
Lean	0,62%	-0,97%	-1,87%	-1,28%	-1,37%	-1,35%	-1,90%

High vibrations



	F1	F2	F3	F4	F5	F6
AmpAv	609	146	119	159	166	119
AmpSD	37,6%	39,1%	60,0%	61,1%	36,4%	31,4%
Shimmer	5,74%	4,62%	5,69%	5,14%	5,21%	5,44%
Lean	-5,66%	4,47%	-6,75%	-5,93%	-6,70%	-2,77%

Fig 4.

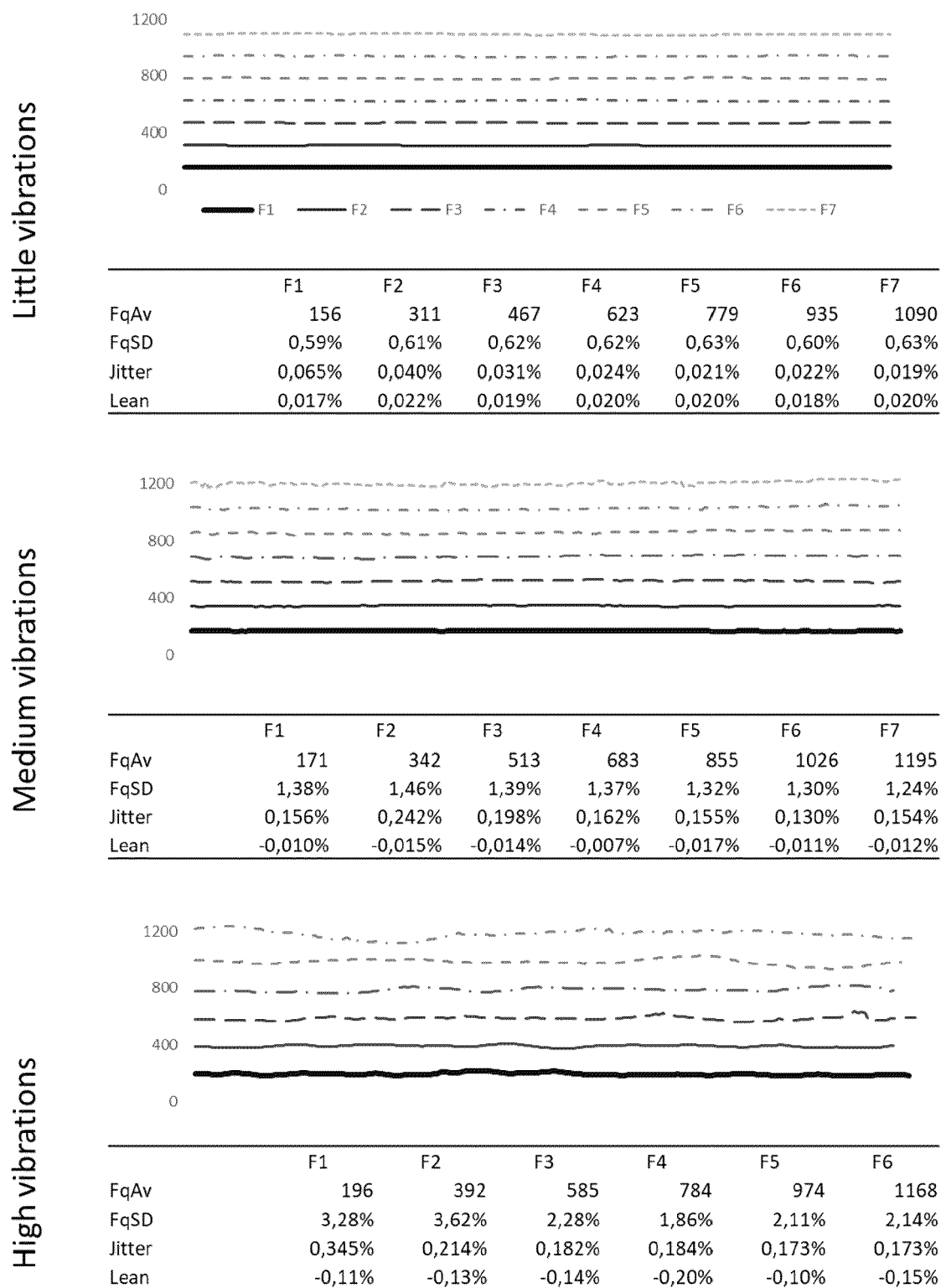


Fig 5.

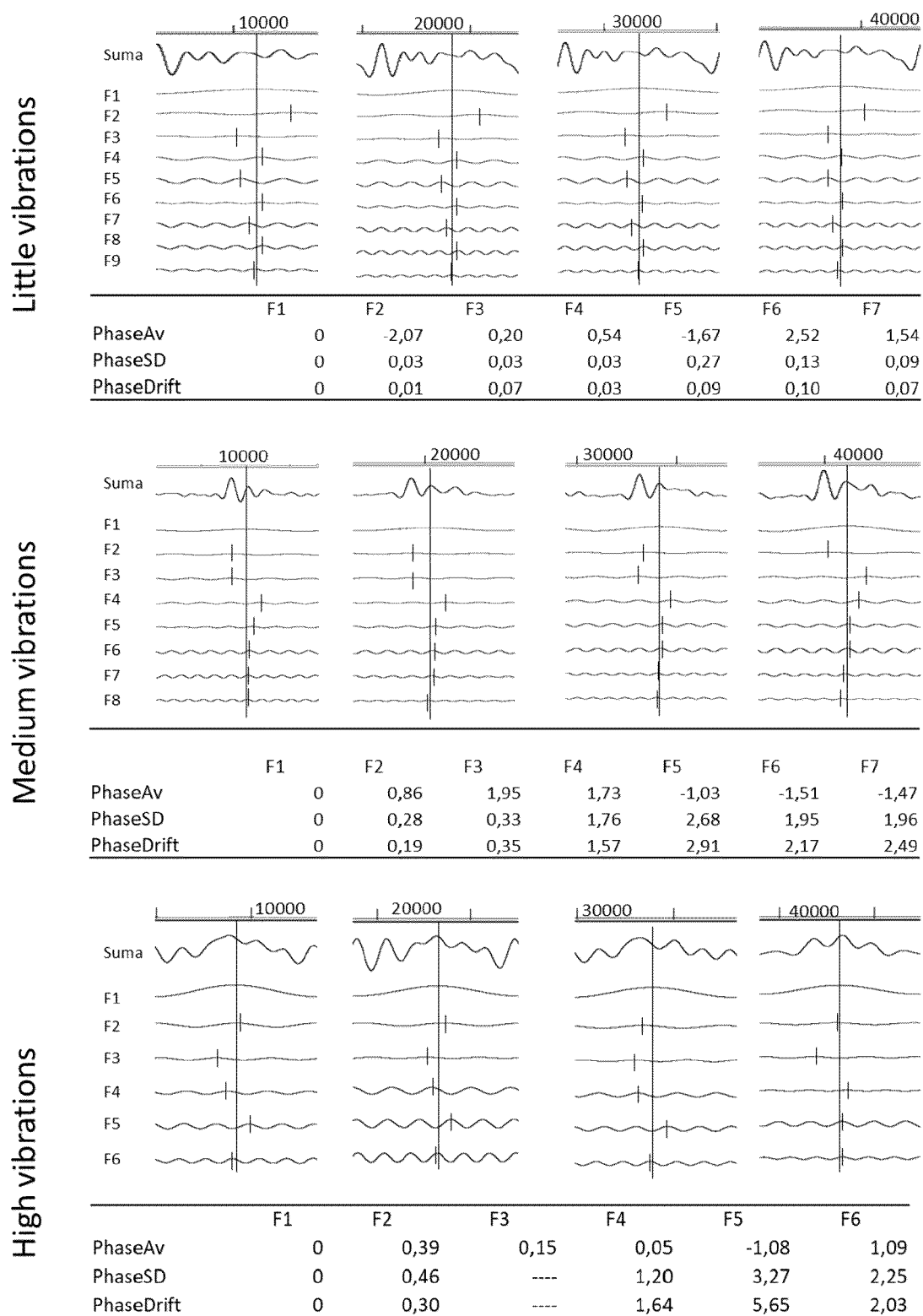
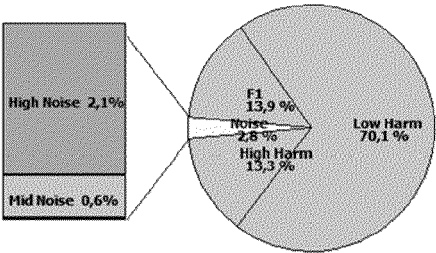
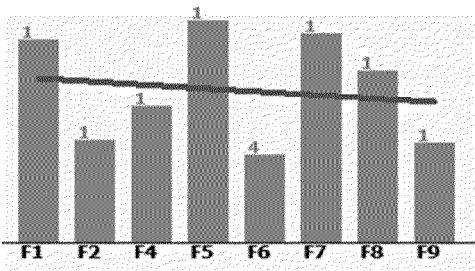


Fig 6.

Little vibrations

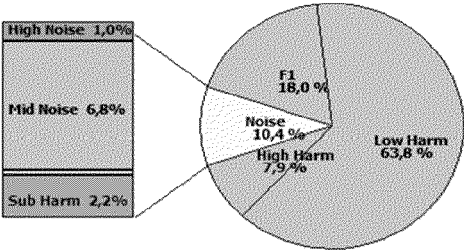


Harm to Noise Ratio	34,9
Harmonic Energy	97,2 %
Noise Energy	2,8 %
Sub Harm Energy	0,0 %

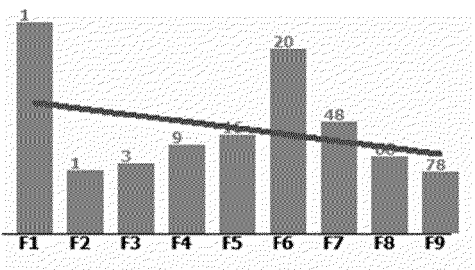


Strong Harm	8
SObd in Strong Harm	11
Average SObd / Harm	1,2
Harmonic Tilt	-0,2%

Medium vibrations

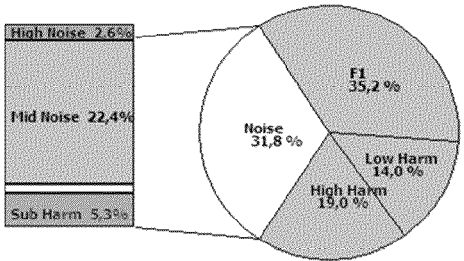


Harm to Noise Ratio	8,6
Harmonic Energy	89,7 %
Noise Energy	7,8 %
Sub Harm Energy	2,2 %

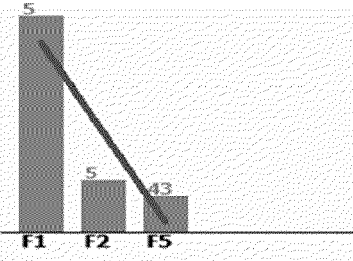


Strong Harm	9
SObd in Strong Harm	244
Average SObd / Harm	27,1
Harmonic Tilt	-0,5%

High vibrations



Harm to Noise Ratio	2,1
Harmonic Energy	68,2 %
Noise Energy	22,6 %
Sub Harm Energy	5,3 %



Strong Harm	3
SObd in Strong Harm	53
Average SObd / Harm	22,7
Harmonic Tilt	-14,7 %

Fig 7.



EUROPEAN SEARCH REPORT

Application Number

EP 23 46 1601

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>PEETERS GEOFFROY ET AL: "The Timbre Toolbox: Extracting audio descriptors from musical signals", THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, AMERICAN INSTITUTE OF PHYSICS, 2 HUNTINGTON QUADRANGLE, MELVILLE, NY 11747, vol. 130, no. 5, 1 November 2011 (2011-11-01), pages 2902-2916, XP012152867, ISSN: 0001-4966, DOI: 10.1121/1.3642604 [retrieved on 2011-11-16]</p> <p>* page 2904, right-hand column, line 43 - page 2915, left-hand column, line 12; figures 3,4; table I *</p> <p>-----</p>	1-15	<p>INV.</p> <p>G10L25/03</p> <p>G10L25/66</p> <p>G10L25/75</p>
X	<p>XIN ZHANG ET AL: "Discriminant Feature Analysis for Music Timbre Recognition and Automatic Indexing", 17 September 2007 (2007-09-17), MINING COMPLEX DATA; [LECTURE NOTES IN COMPUTER SCIENCE], SPRINGER BERLIN HEIDELBERG, BERLIN, HEIDELBERG, PAGE(S) 104 - 115, XP019075844, ISBN: 978-3-540-68415-2</p> <p>* abstract *</p> <p>* page 106, line 11 - page 110, line 23 *</p> <p>-----</p>	1-15	<p>TECHNICAL FIELDS SEARCHED (IPC)</p> <p>G10L</p>
X	<p>XIN ZHANG ET AL: "Analysis of Sound Features for Music Timbre Recognition", 2007 INTERNATIONAL CONFERENCE ON MULTIMEDIA AND UBIQUITOUS ENGINEERING IEEE PISCATAWAY, NJ, USA, IEEE, PISCATAWAY, NJ, USA, 1 April 2007 (2007-04-01), pages 3-8, XP031086496, ISBN: 978-0-7695-2777-2</p> <p>* page 3, left-hand column, line 18 - page 7, right-hand column, line 16 *</p> <p>-----</p> <p>-/--</p>	1-15	
The present search report has been drawn up for all claims			
Place of search		Date of completion of the search	Examiner
Munich		31 October 2023	Dobler, Ervin
CATEGORY OF CITED DOCUMENTS			
<p>X : particularly relevant if taken alone</p> <p>Y : particularly relevant if combined with another document of the same category</p> <p>A : technological background</p> <p>O : non-written disclosure</p> <p>P : intermediate document</p> <p>T : theory or principle underlying the invention</p> <p>E : earlier patent document, but published on, or after the filing date</p> <p>D : document cited in the application</p> <p>L : document cited for other reasons</p> <p>.....</p> <p>& : member of the same patent family, corresponding document</p>			

EPO FORM 1503 03.82 (P04C01)



EUROPEAN SEARCH REPORT

Application Number

EP 23 46 1601

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	ALICJA A WIECZORKOWSKA ET AL: "Identification of a dominating instrument in polytimbral same-pitch mixes using SVM classifiers with non-linear kernel", JOURNAL OF INTELLIGENT INFORMATION SYSTEMS, KLUWER ACADEMIC PUBLISHERS, BO, vol. 34, no. 3, 26 August 2009 (2009-08-26), pages 275-303, XP019790648, ISSN: 1573-7675 * page 275, line 1 - page 287, line 14; figure 1 * -----	1-15	
			TECHNICAL FIELDS SEARCHED (IPC)
The present search report has been drawn up for all claims			
Place of search		Date of completion of the search	Examiner
Munich		31 October 2023	Dobler, Ervin
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)

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

- US 15686057 A [0003]
- CN 202011522343 A [0004]
- CN 202110775311 A [0006]
- CN 202111677396 A [0007]
- EP 16741938 A [0008]
- EP 3304549 A1 [0033] [0034] [0040] [0057]