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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 December 2012 (13.12.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/168740 A1

(51) International Patent Classification:
G10H 1/00 (2006.01) *G06F 17/30* (2006.01)

(21) International Application Number:
PCT/GB2012/051314

(22) International Filing Date:
11 June 2012 (11.06.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1109731.8 10 June 2011 (10.06.2011) GB

(71) Applicant *(for all designated States except US)*: **X-SYSTEM LIMITED** [GB/GB]; Burchetts Brook, Holmbury Lane, Holmbury St. Mary, Surrey RH5 6NA (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **OSBORNE, Nigel** [GB/GB]; Burchetts Brook, Holmbury Lane, Holmbury St. Mary, Surrey RH5 6NA (GB). **ASHCROFT, Robert** [GB/GB]; Burchetts Brook, Holmbury Lane, Holmbury St. Mary, Surrey RH5 6NA (GB). **ROBERTSON, Paul** [GB/GB]; Burchetts Brook, Holmbury Lane, Holmbury St. Mary, Surrey RH5 6NA (GB). **KINGSLEY, Peter** [GB/GB]; Burchetts Brook, Holmbury Lane, Holmbury St. Mary, Surrey RH5 6NA (GB).

(74) Agent: **ORIGIN LIMITED**; Twisden Works, Twisden Road, London NW5 1DN (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(54) **Title**: METHOD AND SYSTEM FOR ANALYSING SOUND

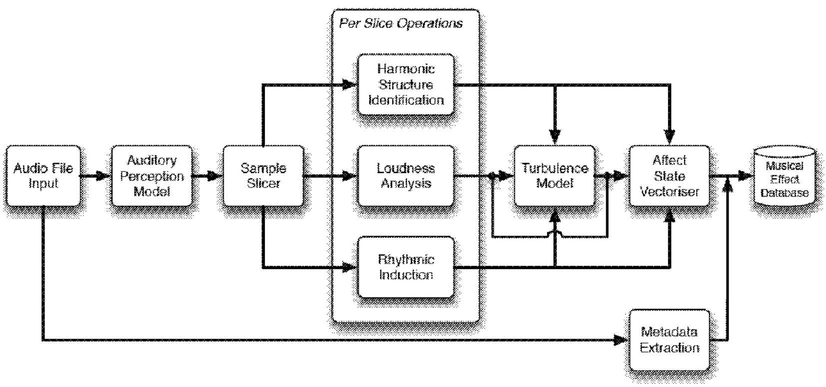


FIGURE 7B

(57) **Abstract**: The present invention relates to a method and system for analysing audio (eg. music) tracks. A predictive model of the neuro-physiological functioning and response to sounds by one or more of the human lower cortical, limbic and subcortical regions in the brain is described. Sounds are analysed so that appropriate sounds can be selected and played to a listener in order to stimulate and/or manipulate neuro-physiological arousal in that listener. The method and system are particularly applicable to applications harnessing a biofeedback resource.

WO 2012/168740 A1

METHOD AND SYSTEM FOR ANALYSING SOUND

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to a method and system for analysing sound (e.g. music tracks). Tracks from a database of sounds, for example music, can be analysed in order to predict automatically the effect or impact those sounds will have on a listener.

2. Technical Background

It is well established that there are specific levels of neuro-physiological arousal (related to mood, states of mind and affect) best suited to particular activities such as study, relaxation, sleep or athletic performance. However, because these levels of arousal result from complex interactions between the conscious mind, environmental stimuli, the autonomic nervous system, endocrine activity, neurotransmission and basal metabolism, it is difficult to control and sustain them.

It is also well established that there is a universal human response to music based on a complex set of functions ranging from perceptual systems, by way of cerebral cortex and other processing, to activation of core emotional centres of the brain and the somatic systems. It is similarly well established that these functions reside in parts of the brain such as, for example, the cochlea, primary auditory cortex, pre-motor cortex, amygdala and the periaqueductal grey (and so on). Rhythm, for example, has a measurable effect on the pre-motor cortex, autonomic nervous system, somatic systems, the endocrine system and neurotransmission. Other aspects of musical structure and experience may also influence human neurophysiology, as described below.

3. Discussion of Related Art

Three ways are known of analysing music for arousal and counter-arousal using humans (for brevity, the term ‘arousal’ will at times be used to include counter-arousal in this document). The first method entails the judgment of an individual, who might be either an expert or the subject him or herself. The second method is by testing many people and asking them how they feel in response to different music tracks. Neither is reliable because each is too subjective.

The third method is to analyse metrics computed as a function of the music itself (usually tempo, but may also include a measure of average energy), and relate such metrics to the desired state of arousal of the subject. There are several such systems, some of which are cited below. Most rely on either 'entrainment' (in the Huygens sense, namely the tendency to synchronise to an external beat or rhythm) or on the association of increased tempo (and in one known case, energy) with increased effort or arousal (and the converse for reduced tempo and energy).

Examples of prior art systems that use music selected according to tempo to manipulate arousal and counter-arousal include US 282045, US191037, US113725, US 270667, WO 151116, US 5267942). This art may use beats per minute as calculated to predict entrainment or may, as in US 060446, modulate tempo in order to improve entrainment. Although this art may be directionally correct, and by extension of Huygens' entrainment principle, it is likely to work to some extent with some repertoire, tempo is both difficult to detect automatically and on its own may best be used to calculate neuro-physiological effect in the limited circumstances where the tempo is both easily and accurately detected and where it is close to the current heart rate of the listener (see next paragraph). Any significant divergence and the entrainment effect is likely to be lost. Most significantly, as discussed below, effective rhythmic entrainment depends on more than beats per minute, and is inseparably synergetic with and dependent on other musical generators of arousal, such as, for example harmonicity and turbulence.

US 5667470 relies on the fulfilment or denial of expected outcomes in music in comparison with established patterns in the repertoire, while US 4883067 introduces the concept of training the brain to replicate positive patterns of neurological activity by association with certain sound signals. One patent, US 5267942, cites the iso-moodic principle documented by Altshuler in 1948 as evidence for its assertion that for the tempo of music to have any effect in entraining heart rate it must lie within the 'entrainment range' of the individual's actual heart rate, i.e. close to it. This introduces the notion that the neuro-physiological effect of a piece of music depends on the initial state of the subject, which means that the effect of any given piece of music is relative rather than absolute. Reference may also be made to US 2007/0270667 attempts to use biometric feedback to manipulate arousal.

Reference may also be made to psychoacoustics. Psychoacoustics has been extensively used in music compression technology (e.g. MP3), but another application is documented in US 7081579, which describes an approach to song similarity analysis based on seven measured characteristics: brightness, bandwidth, volume, tempo, rhythm, low frequency noise and octave. These techniques can identify 'soundalike' music (of which there is much these days) but cannot be used to predict the effect of music in neuro-physiological terms.

SUMMARY OF THE INVENTION

The invention is a computer implemented system for analysing sounds, such as audio tracks, the system automatically analysing sounds according to musical parameters derived from or associated with a predictive model of the neuro-physiological functioning and response to sounds by one or more of the human lower cortical, limbic and subcortical regions in the brain;

and in which the system analyses sounds so that appropriate sounds can be selected and played to a listener in order to stimulate and/or manipulate neuro-physiological arousal in that listener.

The model is a 'predictive model of human neuro-physiological functioning and response' because it predicts how the brain (e.g. structures in the lower cortical, limbic and subcortical regions, including the related autonomic nervous system, endocrine systems, and neuro-transmission systems), will respond to specific sounds.

In one implementation, tracks from a database of music are analysed in order to predict automatically the neuro-physiological effect or impact those sounds will have on a listener. Different audio tracks and their optimal playing order can then be selected to manipulate neuro-physiological arousal, state of mind and/or affect – for example to move towards, to reach or to maintain a desired state of arousal or counter-arousal, state of mind or affect (the term 'affect' is used in the psychological sense of an emotion, mood or state).

We can contrast this system with conventional psychoacoustics (underlying for example MPEG MP3 audio compression algorithms) because psychoacoustics in general deals with how incoming pressure waves are processed by modelling the signal processing undertaken by, for example, the cochlea and primary auditory cortex, whereas the present invention deals with the effect of sound – e.g. the neuro-physiological functioning and response to sound in the lower cortical, limbic and subcortical regions of the brain. Also, the science of psychoacoustics is not concerned with selecting specific sounds for the purpose of stimulating and manipulating desired states of arousal in a listener.

We can also contrast this system with a trivial model of musical effect, such as increased tempo leads to greater arousal. Missing entirely from such model is a generalised understanding of neuro-physiological functioning and response to sound; furthermore, in practice, such a model is so weak as to have no genuine predictive property and, for the reasons given above, is not a general solution to the technical problem of selecting different sounds so as to stimulate and manipulate arousal levels in a listener, unlike the present invention.

The musical parameters derived from or associated with the predictive model may relate to rhythmicity, and harmonicity and may also relate to turbulence – terms that will be explained in detail below. The invention may be used for the search, selection, ordering (i.e. sequencing), use, promotion, purchase and sale of music. It may further be used to select, modify, order or design non-musical sounds to have a desired neuro-physiological effect in the listener, or to permit selection, for example in designing or modifying engine exhaust notes, film soundtracks, industrial noise and other audio sources.

The invention is implemented in a system called **X-System**. X-System includes a database of music tracks that have been analysed according to musical parameters derived from or associated with a predictive model of human neuro-physiological functioning and response to those audio tracks. X-System may include also a sensor, a musical selection algorithms/playlist calculator for selecting suitable tracks and a connection to a music player. Once the sensor is activated, the system diagnoses the subject's initial level of neuro-physiological arousal and automatically constructs a playlist derived from a search of an X-System encoded musical or sound database that will first correspond to or mirror this level of arousal, then lead the listener towards, and help to maintain her/him at, the desired level of arousal. The playlist is recalculated as necessary based on periodic measurements of neuro-physiological or other indicative signals.

Measurement of neuro-physiological state may be done using a variety of techniques, such as electro-encephalography, positron emission tomography, plasma, saliva or other cell sampling, galvanic skin conductance, heart rate and many others, while prediction of response may be achieved via any suitable set of algorithms that are first hypothesised and then refined through testing. Any given set of algorithms will be dependent on the stimulus being modelled and the biometric by which the effect of the stimulus is to be measured, but, even given constant parameters, there are a number of valid mathematical approaches: the specific algorithms we describe in this specification themselves are therefore not the most fundamental feature of the invention, even though most algorithms in the system are unique in conception and implementation. Nor are the particular biometrics chosen to measure neuro-physiological state, though galvanic skin conductance and heart rate are both suitable for general use because they enable measurements to be taken easily and non-invasively, while both give a good indication of arousal or counter-arousal in the autonomic nervous system, which is in turn largely synergetic with endocrine activity and related neurotransmission.

X-System represents an improvement upon existing art in that it: a) describes the bio-active components of music (beyond tempo and energy) by reference to the brain's processing of audio stimuli, including music, and b) describes how any given sound source may be calibrated to the initial state of the subject in order to have the maximum entrainment effect. It offers the advantage over many other systems that it requires neither the modulation of tempo (tempo modulation is known from US 2007/0113725, US 2007 0060446 A1, US 2006/0107822 A1) nor the composition of psycho-acoustically correct, synthetic music (known from US 4883067) to achieve its effect. X-System offers the possibility of harnessing the entire world repertoire of music to the modulation of affect without needing to manipulate the rendering of the music in any way.

X-System is based on a paradigm we shall refer to as the 'Innate Neuro-physiological Response to Music' (INRM - we will describe this in more detail below), and a unique informatic modelling of one or more of lower cortical, limbic and subcortical functions related to these responses. X-System has a unique capacity to analyse music tracks automatically and establish the potential to generate levels of arousal and counter-arousal in the listener. This unique method of analysis is a human universal and may be applied to music of all human cultures as well as to environmental and other sound sources. X-System is capable of categorising databases of music and sound according to core emotional effect. X-System may implement automatic categorisation remotely, for example for personal repertoires. X-System may also have the capacity to detect the state of mind and body of the user, using a unique radio electrode and microphone based conductance/heart rate sensor and other devices. X-System may use this sensor data to sub-select music from any chosen repertoire, either by individual track or entrained sequences, that when listened to, will help the user to achieve a target state of excitement, relaxation, concentration, alertness, heightened potential for physical activity etc. This is achieved by analysing music tracks in the user's database of music (using the musical parameters derived from the predictive model of human neuro-physiological response) and then automatically constructing a playlist of music, which may also be dynamically recalculated based on real-time bio-feedback, to be played to the user in order to lead her/him towards, and help to maintain her/him at, the desired target state.

As noted above, X-System models the effect of music on specific parts of the lower and middle brain, including the limbic system and subcortical systems, but these are not the only parts of the brain that respond to music. Other centres govern a more personal experience involving preference, culture, memory and association, the meaning of the lyrics, the historical context in which they were written, the knowledge of the circumstances of the performer or composer and other factors. These too have a significant effect, so it is important not to expect any piece of music to have an absolute effect on any one individual. INRM describes an important part of, but not all, musical effect. A prediction that certain pieces of music will calm the listener, or even induce sleep, is not like a drug or an anaesthetic, where the effect of a certain dose can be predicted with reasonable accuracy and where that effect cannot be resisted by conscious effort. Nevertheless, tests confirm that each of the elements of the brain that the INRM model is based on are strongly linked to arousal and counter-arousal. Music though, has its greatest effect when selected appropriately to accompany a desired state or activity and X-System offers an automated means of selecting music that is always appropriate to what the listener is doing, which can be very effective in a host of situations from treating anxiety to enhancing relaxation or concentration, or stimulating creative 'flow', or in bringing power and fluency to athletic activity. The brain modelling that underpins X-System offers a further capacity offered by no other existing categorisation system: it is **universal**; X-System may accurately predict levels of physiological arousal for all music of the world repertoire, whether it be Western classical and pop, Chinese or Indian classical or folk music, African pop or roots, or avant-garde electronica or jazz.

X-System has proven to be capable of outperforming expert musicologists in predicting, over a broad repertoire, a general index of arousal/counter-arousal based on the biometric parameters of heart rate and galvanic skin resistance, but were these biometric parameters to be different the equations, which we will describe later in this document, would almost certainly need to be modified; equally, there are many mathematical techniques familiar to those skilled in the art that could have been used to predict the neuro-physiological effect of a piece of music and any one of many might produce equally satisfactory results. A key feature of this invention therefore lies in the identification of the patterns in music that are neurophysiologically active ('bio-active') and that may have a predictable effect on human neurophysiology, including arousal and counter-arousal.

Other aspects of the invention

We list fifteen further aspects of the invention below, each of which may also be combined with any other:

1. A computer-implemented method of categorizing sound (such as any piece of music regardless of genre or cultural origin) (e.g. according to musical parameters derived from a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to the pieces of music) in such a way that it may be selected (e.g. automatically based on biometric data captured by a sensor) to entrain neuro-physiological arousal towards a target level; this may occur while directing the listener towards one or more among a number of pre-assigned states of mind and/or affect, or in order to direct the listener towards one or more among a number of pre-assigned states of mind and/or affect.

2. Automatic categorisation of sound (such as pieces of music) in a remote database (e.g. according to musical parameters derived from a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to the pieces of music). This includes the idea that we can search/discover music that has similar X-System deep structures and cross match conventional categorisation schemes (Gracenote® etc) to X-System. As an alternative to, or in addition to, automatic categorisation, X-System provides selection and 'push' for commercial or promotional purposes, or a methodology for description or detection of particular music, for all applications, not only entrainment. An example is a computer-implemented method of categorizing any piece of music regardless of genre or cultural origin according to its Innate Neuro-physiological Response to Music for the purpose of search, navigation, music discovery, retrieval and selection.

We now expand on the concept of search/discovery, in which X-System provides for automated search of musical remote or local databases and of X-System encoded services. In this application, users may:

- Search for music that has similar signatures to the music they tag that they like, by pressing a 'find more' or 'I like' key on their computer or Smartphone X-System device App. This will cross-match X-System encoding of universal arousal information with other individual features within an App (such as favourites, or frequently listened to) in order to create a new level of personalisation;
- Search by and for patterns of listening preferences amongst social network groups, such that by sharing my preferences and choices and communicating them to my friends, they will see the relationships between my emotional response to particular tracks and comparisons with others in the network;
- Search by musical or experiential journey, such that a particular sequence of music can be stored, for example, on my Smartphone and repeated when I press 'I liked that sequence, store it so I can play it again';
- Search by finding patterns and relationships between tracks users tag as 'I like', such that similar combinations of say genre, musician, activity and X-System encoded arousal data can drive recommendations. So, for example, X-System will generate a playlist suggestion that will combine jazz, particular Miles Davis tracks, writing an essay, concentration and arousal levels, if a similar combination has been tagged from an earlier listening sequence (the tagging of activity being part of the Smartphone App); and
- Search on Google and other web sites for X-System encoded information, such that, for example, music, video or other web content is categorised and tagged, either automatically; or in collaboration with search engine providers such that it 'advertises' X-System arousal or mood states; or according to visitors who tag web sites automatically as they view pages.

3. An automated diagnosis of the level of lower cortical, limbic and subcortical neuro-physiological arousal of an individual and expressing it as a value in order to correspond to the musical effect of any one of a theoretically unlimited number of pieces

of music in a database. Alternatively or additionally, there may be provided a method of trial and error of self-diagnosis e.g. by song selection as described above.

4. A computer-implemented method of creating a playlist of tracks generated by automatically (or indeed manually) analysing musical parameters derived from a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to the pieces of music in order to entrain arousal and direct state of mind and/or affect. Optionally, this may include:

a) choosing a subset of the music in a database by reference to existing descriptive metadata, if available, such as genre or user-created playlist; b) selecting from this subset of music a number of pieces that will correspond to the user's initial level of lower cortical, limbic and subcortical neuro-physiological arousal by matching it to music contained in the relevant row of the musical effect matrix (we will explain this matrix in more detail later); c) selecting a target state of mind and/or affect; d) selecting a series of ascending or descending musical effect values which correspond to the expected entrainment path from the initial to the required level of neuro-physiological arousal; e) on the basis of this series of values, selecting qualified content from the music database; f) choosing at random a playlist from the qualified content subject to other rules such as genre preference, the anti-repetition rule (see 'Musical Selection Algorithms' below) or the Unites States' Digital Millennium Copyright Act (DMCA) rules; g) repeating the calculation of the playlist at intervals, based on continual biometric feedback - for example, the playlist may be recalculated once per minute, based on biometric feedback including the most recent feedback.

5. A method of determining the sufficiency of a (e.g. personal) database of music for the entrainment of affect and of then displaying information to the user with regard to sufficiency or insufficiency.

6. A method of recommending a complement of musical content for a personal database of music in order to ensure sufficiency, by using musical parameters derived from a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to that music.

7. A method of selecting music which has a similar musical effect, (e.g. according to musical parameters derived from a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to the pieces of music). This may include a search by X System code.

8. A method of categorising music according to its musical effect rather than its descriptive attributes.

9. A method of ordering a series of pieces of music in a playlist by matching the musical effect of each piece with a temporal series of values described by a musical effect vector.

10. A method of manipulating the arousal of a user by using any of the above methods or systems.

11. A method to modify the properties of ambient sound in any given environment, in order to produce a desired neuro-physiological response in the listener, by using any of the above methods or systems. And the use of this as a selection, control or design tool to define such responses.

12. A system adapted to perform any of the above methods.

13. Software (whether device-resident, network resident or elsewhere), firmware, SoCs or audio stacks programmed or adapted to perform any of the above methods or to form part of the system described above.

14. A computing device, such as a smartphone or tablet, adapted to manipulate the arousal of a user by using any of the above methods or by using or including any of the above systems, software, firmware, SoCs or audio stacks.

15. Sensors adapted to work with the computing device defined above.

Some more generalised observations now follow:

It is the identification of which structural and experiential phenomena in music activate which parts of the primitive brain, the development of techniques to measure them using digital signature analysis and the construct of a series of generic models that use relatively simple equations to predict levels of activation of relevant regions and organs of the brain, and in turn their effect on biometric indices, that are some of the key aspects of this invention.

Examples of the present invention may work with all musical genres and do not depend upon there being any pre-existing metadata in a database of digitised music. The database may be assembled by the user from his or her own collection and stored on a local playback device, in which case the music on the database may be profiled remotely, it may be supplied pre-analysed on a digital storage device, or it may be streamed from a central server. In these latter cases, the music may be associated with other data and/or digital media in order to enhance the user experience, or signature excerpts may be profiled and included in order to accelerate the desired effect.

The invention may be implemented as application software on either a remote server, on the music playback device itself or on another device that is connected to the music playback device either directly or via either a local or wide area network, or firmware or embedded in a chip; it may form part of an audio stack or may be used as part of a set of design tools. These implementations may enable real-time analysis of music tracks and other sounds, all done locally within a portable computing device such as a smartphone or tablet, or remotely on a server, or some combination of distributed local and server based processing. All such deployments will also support a consistent API to enable application vendors and service providers to access system capability, for example, to enable new application to be constructed and deployed.

If the necessary metadata are available, a preferred musical style may be chosen among those on the music database; if not, the system may select from the whole music database rather than a chosen subset.

The following terms are taken to have specific meanings in this document:

'Level of neuro-physiological arousal': an index calculated, for example, as a function of galvanic skin conductivity and pulse rate, though other parameters may also be selected including where more complex measurement is required. Different levels of neuro-physiological arousal facilitate different activities, states of mind and affect.

'State of mind': the dynamic relationship between functional areas of the brain associated with different types of thought such as creativity, learning, meditation, imagination etc.

'Affect' (noun): as used in psychology to mean feeling or emotion and in psychiatry to mean expressed or observed emotional response. Mood.

'Musical Effect': the state of mind or mood that is provoked by a given piece of music and the influence it has upon neuro-physiological arousal.

'Sound': includes any sound, including music as that term is conventionally understood but also extending to other sounds such as the ambient or background noise in a workplace, cinema, home, shop, vehicle, car, train, aircraft: anywhere where sound can in theory effect listener arousal. For example, tuning car exhaust notes would be one example; modifying engine sounds another. Sounds of nature (wind, ocean etc.), sounds of animals, sonifications (planets, stars, flowers, trees, financial markets, cell activity etc.) are other examples of 'sounds'. In this document, we will refer to 'music', but that term should be expansively construed to include not merely music in the sense of the art form in which voices and/or instruments are combined to give harmony, beauty or self-expression, but also all other forms of sound, as that term is expansively defined above.

A note on terminology: The primary auditory cortex is situated in the temporal lobes of the neo-cortex - the most "evolved" part of the brain, but it is essentially "low" in the system and hence 'lower cortical'. Organs critical to X-System, such as the hippocampus and amygdala are generally described as "limbic" (from the Latin "limen, liminis", meaning "threshold", i.e. at the lower limit of the neo-cortex). These are close to emotion-related areas such as the nucleus accumbens, and periaqueductal grey, sometimes also regarded as limbic. The limbic system may also be described as the archicortex and paleocortex - the "main, initial or ruling" and "old" cortex. Finally, many X-System areas related to rhythm, core emotion and movement are sub-cortical, for example the basal ganglia and cerebellum.

X-System therefore relates primarily to lower cortical, limbic and sub-cortical areas of the brain, concerned with fundamental and universal responses to music, as opposed to more cognitive-related, culture-related and reflective areas of the neo-cortex.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a graphical representation of the neural elements involved in audio processing applicable to the X-System. Elements enclosed within the solid boxes are part of the current model; elements contained in the dashed boxes may be included in the model.

Figure 2 shows an overall system construction where the user of the system both selects their desired affect goal and is the recipient of the generated output.

Figure 3 shows an overall system construction where selection of target affect is made by a party external to the user of the system.

Figure 4 shows an implementation of the X-System invention where all aspects of the software reside on the users PC (the term ‘PC’ should be construed expansively to cover any computing device of any form factor, including any device capable of performing computing functions).

Figure 5 shows an implementation of the X-System invention where a primary music library, and analysis software resides on a user PC, with the ability to transfer a selection of music to a personal music player device, which then generates a dynamic playlist based on the available music.

Figure 6 shows an implementation of the X-System invention where an external service provider offers an analysis tool via a network connection. Audio may reside on either the user’s PC or be "streamed" by the service provider, and a database of stored musical affect may be used to minimise track analysis.

Figure 7A is a detailed block diagram showing the major components of the X-System audio analysis tool used in analysing harmonicity.

Figure 7B is a detailed block-diagram showing all of the major components of the X-System audio analysis tool.

Figure 8 is a detailed block-diagram showing the major components of the X-System music playback and monitoring application.

Figure 9 shows schematically arousal as a function of time, for Excite, Maintain or Relax pathways.

Figure 10: Modelling of the cochlea and primary auditory pathways is achieved through the use of an A-weighting filter. This attenuates lower frequencies and amplifies higher

frequencies, dropping off again quickly towards the upper frequency limit of human hearing.

Figure 11 shows Beat Energy as a function of time.

Figure 12 shows Harmonic Energy as a function of time.

Figure 13 shows Harmonic Cost as a function of time.

Figure 14 shows Volume as a function of time.

Figure 15 shows Harmonic Energy as a function of time.

Figure 16 shows sample categorisation from the Miles Davis repertoire.

Figure 17 shows an example of other manual categorisations, in which tracks are further sorted into stable, rising and falling vectors.

Figure 18 shows an example in which movements from Beethoven symphonies have been categorized according to the vectors.

DETAILED DESCRIPTION

This Detailed Description has the following sections:

A. High Level Concepts

There is scientific evidence that music entrains and shapes arousal, state of mind and affect through direct neuro-physiological engagement; this invention concerns the discovery and general method of determination of the Innate Neuro-physiological Response to Music, and includes a novel method of harnessing this phenomenon. As noted above, this invention is implemented in a product called X-System. X-System harnesses the potential of music to effect neuro-physiological changes in listeners, in particular in relation to arousal and counter-arousal and associated states of mind, working at the level of the most fundamental, innate, neuro-physiological functioning and response of the limbic, lower cortical and sub-cortical regions of the brain.

It differs from other approaches to music categorization in that it is not concerned with musical similarity, either by semiotic labelling or the analysis of acoustic characteristics. It also differs from standard therapeutic approaches, such as classification of mood.

X-System works through predictive, deterministic modelling of INRM (Innate Neuro-physiological Responses to Music) (Osborne 2009, unpublished), see **Figure 1**, and the structuring of pathways towards target states of body and mind. **Section B** explains INRM in more detail. In brief, the INRM paradigm assumes a standard interpretation of audition, from the auditory canal to the oval window of the cochlea. The cochlea itself is modelled to reproduce the characteristics of human audition. The paradigm further assumes neural pathways to the inferior collicus and primary auditory cortex. Levels of arousal related to pulse and rhythmicity are predicted through a simple modelling of mirror neuron and pre-motor related systems, including tempo induction and indices of rhythmic power and density. Other bio-active characteristics of music may also be modelled such as the identification of rhythmic patterns in the right anterior secondary auditory cortex, among others.

X-System additionally models the functioning of Heschls gyrus, the posterior planum temporale, superior temporal sulcus and circular insular sulcus to predict arousal-related qualities of timbre and exponential series-related frequency structures, including octave equivalences. There are other modelling possibilities such as arousal-related effects among chroma (individual notes of melodies) in the planum polare using, for example, harmonicity indices.

Finally, general levels of 'turbulence' are calculated as a prediction of arousal and counter-arousal in core emotional locations and organs such as the periaqueductal grey and amygdala.

The predictive arousal and counter-arousal values calculated are combined to model the process of arousal and counter-arousal in the autonomic nervous system, and associated systems such as the HPA (hypothalamic-pituitary-adrenal) axis.

A sensor may optionally be used to establish the state of arousal of the user, and music categorised by predictive modelling of the INRM paradigm can then be streamed/played back to achieve the target arousal state for that user. In an alternative implementation sensors are not provided. Instead, both initial and target states are self-selected, either directly or indirectly (such as, for example, by selecting a 'start song' which has an arousal value relative to the user's true current state). For example, where the user makes a poor initial selection, he/she might skip from song to song initially until one is found (i.e. by trial and error) that is both 'liked' and 'fits' with their initial state. From there, X-System, in a sensor-less implementation, may create a playlist tending towards the desired arousal state based on expected normal human response.

In another alternative, an implementation is provided for a group of people as a system with software but no sensor, reliant on average expected response. An application is for 'crowd' applications, where an automated disc jockey (DJ) would be able to manipulate the mood of a crowd at a party.

Other alternatives include applications controlling the personal audio environment by sending emotional cues to the system via sensors, and polling group emotion via either sensor or sensorless inputs, in order to entrain the person or group towards a desired response.

Other alternative applications include the search, selection, description, detection, sharing or promotion, of music based on its neuro-physiological content.

As in the case of all systems and activities related to music and arousal, there are variations in response among individuals, and variations as a result of extreme or unusual states of body and mind, medication etc. The strength of X-System is that it works on the basis of the most fundamental physiological responses, which may act in an ethical and democratic synergy with conscious and unconscious consent of the user. A further strength of the INMR-based categorisation system is that it may be applied to the music of any human culture, and indeed both to sound design and sounds of the natural world.

B. The Innate Neuro-physiological Response to Music (INRM) in more detail

Figure 1 shows a simplified model of the neural structures related to auditory processing and interpretation. The X-System example of the invention may model the functioning or behaviour of these systems in response to sound (e.g. musical) stimulus as described in the following sections.

The Innate Neuro-physiological Response to Music Paradigm is a predictive, deterministic model of the mind and body's most fundamental response to music. Although responses to music are profoundly influenced by culture, personal history and context, there are basic neuro-physiological reactions that are universal to all musical experience. A substantial body of recent research in neuro-physiology and neuroscience, including evidence from functional Magnetic Resonance Imaging, EEG and Positron Emission Tomography, as well as studies related to endocrine and autonomic activity has made it possible to build a predictive model of how the lower cortical, sub-cortical and limbic parts of the brain react to sound.

X-System makes use of the following protocols for audio input. Input is taken from uncompressed WAV files or any other suitable format (X-System can use lower quality file formats when undertaking remote categorisation – e.g. categorising music tracks on a remotely held server or personal device. Equally, higher quality file formats may be more appropriate in other circumstances). If the track is in stereo, we combine both channels by averaging them. This is particularly important, for example, for 1960s tracks, where some loud instruments were positioned full left or right. This should not cause interference unless the audio has passed through faulty stereo equipment (e.g. a misaligned tape head). The track is split into sections of a given length, and the analysis is carried out independently for each section.

Figure 7A is a block diagram showing the major components in X-System for analysing harmonicity and **Figure 7B** is a block diagram representation of all of the major components of the musical analysis tool. The operation of the major components will be described in the remainder of this Section B.

B.1 The cochlea and primary auditory pathways

Modelling of the cochlea and primary auditory pathways is achieved through the use of an A-weighting filter, as specified in IEC 61672. This attenuates lower frequencies and amplifies higher frequencies, dropping off again quickly towards the upper frequency limit of human hearing; the filter 'knee' is at around 6kHz. This weighting is required to ensure that (as in human audition) high energy lower frequency sounds do not overwhelm other spectral information. See **Figure 10**.

B.2 Harmonicity: Heschl's Gyrus and associated tonotopic maps

"Harmonicity" describes the correspondence of sound (e.g. music) to the pattern of the harmonic series (harmonic series are present in the sound you hear when the winds blows through a hollow tree, run your finger lightly up the string of a violin or guitar, or blow progressively harder on a single note on a flute). The harmonic series is a universal pattern of concentrations of sound energy in symmetrical resonating objects: a fundamental tone f, sounds together with its harmonics f2, f3, f4 etc. This pattern has been important throughout the evolution of sentient life forms, from the harmonic resonance of the primal cell, through the perceived "safety" of harmonic sounds in the environment, to the pleasing harmonic resonances of musical instruments and the human voice. "Harmonicity" or correspondence to the pattern of the harmonic series is detected by Heschl's Gyrus, located in the primary auditory cortex of the brain. Harmonicity activates centres of counterarousal and pleasure in core emotional centres of the brain. Inharmonicity, or lack of correspondence to the harmonic series activates systems of arousal.

X-System models the functioning and response of Heschl's Gyrus to sound by determining levels of harmonicity and inharmonicity. This may be a complex process. Musical structures may involve several fundamentals each with their own harmonic or inharmonic spectrum.

X-System is unprecedented in that it combines all emotional processing of pitch and timbre in two harmonicity-related algorithms. Timbre (the internal structure "colour" of a sound), harmonicity (the extent to which the internal structure corresponds to the pattern of the harmonic series) and individual pitches are initially processed in the primary auditory cortex. The main area for processing timbre is the posterior Heschl's gyrus and superior temporal sulcus, extending into the circular insular sulcus (McAdams et al 1995; Griffiths et al 1998; Menon et al 2002). Pitch is processed progressively deeper in areas surrounding Heschl's gyrus: chroma (or differences of pitch within the octave, as in most conventional melodies), activate bilateral areas in front of Heschl's gyrus and the planum temporale, while changes in pitch height (octave transpositions and the like, as in the difference between a man and woman singing the same tune) activate bilateral areas in the posterior planum temporale (Brugge 1985; Pantev et al 1988; Recanzone et al 1993; Zatorre et al 1994; Warren et al 2000; Patterson et al 2002; Formisano 2003; Decety and Chaminade 2003; Jeannerod 2004;Talavage 2004). Harmonicity and pitch structures activate areas of the amygdala and hippocampus, and in turn the autonomic nervous system, core emotional centres, and endocrine and neurotransmission systems (Wieser and Mazzola 1986; Blood and Zatorre 2001; Brown et al 2004; Baumgartner et al 2006; Koelsch et al 2006). X-System predictively models the neurophysiological sensing of simple timbre (Heschl's gyrus, superior temporal sulcus, circular insular sulcus) by analysing windows of vertical harmonicity: X-System detects a principal fundamental through calculation of the harmonic product spectrum, then establishes degrees of harmonicity both within and among the spectra of different fundamentals. This analysis is applied both "vertically" to instantaneous moments, and "horizontally" to progressions of pitches and spectra in time (related to the tonotopic mapping of the area around Heschl's Gyrus) and expressed in terms of linear harmonic cost.

In one very simple implementation, the mean values of linear harmonic cost (C) and instantaneous harmonicity (H) are combined to calculate the inharmonicity (I) of a piece where:

$I = C/10 - H$

This equation is a non-limiting example of how inharmonicity can be calculated and other ways of linking I to C and H may well be appropriate; furthermore, I may be defined in terms of other or additional variables, as may C and H. See **Figures 12** and **13**, showing harmonic energy and cost as a function of time.

More details on Harmonicity calculation now follow:

B.2.1 Spectral analysis

First the STFT of the audio is taken with a window length of 8192 samples and an interval of 2250 samples (0.05 seconds). This produces a 2D array of time vs frequency.

B.2.2 Cochlear modelling

As in the case of rhythmic processing, analyses are performed on a transformed instance of the input sample data, which accounts for certain aspects of the auditory pathway, primarily the cochlea pick-up. The behaviour of the cochlea is well understood and accurate models have been developed. We apply a frequency-dependent gain function to the input signal, which attenuates bass signals and amplifies treble components, with a filter "knee" at around 6kHz. The exact transform used is the "A Weighting" as specified in IEC 61672.

B.2.3 Fundamental frequency detection

For each time slice of the STFT array, the fundamental frequency is determined using the harmonic product spectrum method, as follows:

- Take the frequency spectrum, and produce copies of it compressed along the frequency axis by factors of 2, 3, 4 and 5.
- Multiply all 5 copies (including the original)
- The fundamental frequency is the maximum value of the resulting spectrum.

B.2.4 Mean harmonicity

For each time slice of the STFT array, the mean harmonicity is the ratio of harmonic energy to the total energy present in the slice. Harmonic energy is energy found in the following harmonics of the fundamental, as well as of 1/2 and 1/4 of the fundamental: [1 2 3 4 5 6 7]. For each of these harmonics, we sum the energy found in the closest STFT bucket, plus 3 buckets on either side.

B.2.5 Linear harmonic cost

Predictions of activity in, and progression through, areas surrounding Heschl's Gyrus (planum temporale, posterior planum temporale) including chroma, octave changes and chord progression etc. are combined in a single operation, described as "linear harmonicity" or "harmonic cost".

This is entirely unprecedented: it analyses all melodic and harmonic progressions in terms of how far each step deviates from the simple ratios of the harmonic series: Linear harmonic cost arises from STFT time slices whose fundamental frequency differs from that of the previous slice. Time slices with no change in fundamental have a cost of zero.

The fundamental frequency is first normalised by rounding it to the nearest musical note value under the A440 tuning, then shifting it to a single octave. The (normalised) fundamental is then compared to the previous one: If they are identical, the cost is zero.
If the new fundamental is one of the following harmonics and sub-harmonics of the previous (normalised) fundamental (1/9 1/7 1/6 1/5 1/3 3 6 7 9) then the cost is defined as equal to the multiplier of the harmonic or divisor of the sub-harmonic. Otherwise the cost is defined as 15.

Linear harmonic cost is expressed in terms of cost per second. The metric therefore represents both the rate at which the fundamental is changing, and the harmonic distance of the changes. Higher numbers indicate a more stimulating effect.

Linear harmonicity activates similar emotional systems to vertical harmonicity (Wieser and Mazzola 1986; Blood and Zatorre 2001; Brown et al 2004; Baumgartner et al 2006; Koelsch et al 2006).

B.2.6 Harmonicity and Valence

Both vertical and linear harmonicity are powerful indices of valence (Fritz 2009), or whether a sound is "positive" or "negative", "pleasing" or "not so pleasing". Linear harmonicity may track the evolution of valence indices over time - the principle is simply the more harmonic, the more positive valence, the less harmonic, the more negative valence.

It is conceivable that the Heschl's gyrus-related equations may be reconstituted with a different mathematical approach. It is highly unlikely that the planum temporale function could be approached in any different way.

B.3 Rhythmicity: Mirror neurons, the auditory and pre-motor cortex

Human responses to musical rhythm involve a complex set of activations of mind and body systems (Osborne 1. 2009; Osborne 2. 2009; Osborne 3. 2012) including perceptual systems, the dorsal cochlear nucleus, inferior collicus and spinal systems (Meloni and Davis 1998; Li et al 1998) the primary and secondary auditory cortices (Peretz and

Kolinsky 1993; Penhune et al 1999), mirror neurons (Rizzolati et al 2001; Gallese 2003; Molnar-Szakacs and Overy 2006; Overy and Molnar-Szakacs 2009), pre-motor and motor cortices, basal ganglia, vestibular system and cerebellum (Zatorre and Peretz 2001; Peretz and Zatorre 2003;Turner and Ioannides 2009;), the autonomic nervous system (Updike and Charles 1987; Iwanaga and Tsukamoto 1997; Byers and Smyth 1997; Cardigan et al 2001; Knight and Rickard 2001; Aragon et al 2002; Mok and Wong 2003; Lee et al 2003; Iwanaga et al 2005), and finally somatic and core emotional systems (Holstege et al 1996; Gerra et al 1998; Panksepp and Trevarthen 2009). Some of these may be related in particular to the firing of mirror neurons capable of regenerating perceived behaviours, vitality affect and energies encoded in the sound and its manner of performance in the mind and body of the listener. Fast rhythms of high energy activate arousal in both the Autonomic Nervous System and endocrine systems such as the HPA axis. Slow rhythms activate counterarousal.

X-System detects a basic, "default" rhythmic pulse in terms of beats per minute. There are often difficulties in establishing metre, but X System approximates the arousal effect of metrical structures by averaging the accumulation of power of rhythmic events over time. The power of a rhythmic event is defined as the ratio of the energy before the beat to the energy after it. In one very simple implementation, the beats per minute value (B) is combined with the mean of the beat strength (S) to produce a value for rhythmicity (R) where:

$$R=\sqrt{B} * S^2$$

This equation is a non-limiting example of how rhythmicity can be calculated and other ways of linking R to B and S may well be appropriate; furthermore, R may be defined in terms of other or additional variables. R, in general, may be a function of B and S, but the optimal relationship will depend on various factors. See **Figure 11**, showing beat energy as a function of time.

More details on Rhythmicity:

B.3.1 Cochlear Modelling

As explained earlier, aural perception of rhythm is predicted through conventional cochlear modelling: Following audio input, all subsequent analyses are performed on a transformed instance of the input sample data which accounts for certain aspects of the auditory pathway, primarily the Cochlea pick-up. The behaviour of the Cochlea is well understood and accurate models have been developed. We apply a frequency-dependent gain function to the input signal, which attenuates bass signals and amplifies treble components, with a filter "knee" at around 6kHz. The exact transform used is the "A Weighting" as specified in IEC 61672.

B.3.2 Rhythmic Induction

The activations of primitive spinal pathways and the pre-motor loop (including basal ganglia, vestibular system, cerebellum etc.), all concerned with primal responses to rhythmic impulses, are predictively modelled by beat induction, using a specifically calibrated onset window.

Rhythmicity is, of course, a parameter that models the basic tempo of the sample, as well as higher order metrical structures within. It is computed by first determining note onsets, using spectral flux peak detection. These onsets are then used to generate and score a large number of metrical structure hypotheses. Candidate hypotheses are generated, filtered, and scored, using the methods of Dixon [Evaluation of the Audio Beat Tracking System BeatRoot, Journal of New Music Research, 36 (1), 39-50, 2007]. In addition to the methods described therein, we extend the process to include the magnitude of the spectral flux surrounding the onset event in order to estimate higher order structure. The hypotheses generated are filtered and scored using the same methods, with the final output comprising an estimate of the fundamental tempo of the sample, a secondary output in which the tempo is weighted according to the predicted metrical structure, in which the more distinct an accented beat is from the base beat, the higher this value. A confidence value is also expressed as the variance of the distribution of these outputs for all beat hypotheses scoring above a given threshold. This confidence value is normalised to permit comparison across samples.

B.3.3 Auto-correlation

Rhythmic pattern recognition and retention (for example in the secondary auditory cortex of the temporal lobes) is predictively modelled by self-similarity/auto-correlation algorithms (e.g. Foote http://207.21.18.5/publications/FXPAL-PR-99-093.pdf.)

First the audio is Hamming-windowed in overlapping steps; the log of the power spectrum for each window is calculated by means of DFTs (discreet Fourier transforms). these coefficients are perceptually weighted through Mel-scaling. Finally a second DFTis applied to create cepstral coefficients. High-order MFCCs (Mel-frequency cepstral coefficients) are discarded, leaving the 12 lower-order MFCCs, forming 13-dimensional feature vectors (12 plus energy) at a 100Hz rate. These data are then subjected to vector autocorrelation, plotted in a two-dimensional window, where both x and y axes plot the unfolding of the track in time. Areas of "brightness", reading upwards, for example, from the first instant of the track on the x axis, indicate points of similarity, and likely metrical structures.

Density of distribution of points is also used in a predictive index of rhythm-induced arousal (the greater the density, the higher the arousal).

B.3.4 Power

Activation of mirror neuron systems, which detect, among other things, the power, trajectory and intentionality of "rhythmic" activity, is predictively modelled through indices of rhythmic power, including computation of volume levels, volume peak density, "troughs", or the absence of energy and dynamic profiles of performance energy.

B.3.5 Volume envelope analysis

The volume envelope is calculated as the RMS of 5ms slices of the amplitude data.

B.3.6 Volume level

This is simply the mean RMS level over the time period.

B.3.7 Volume peak density

Number of volume peaks per slice (usually 10 seconds), as found by the MATLAB findpeaks function with minpeakdistance=100ms, multiplied by the mean height of the peaks above the volume mean, divided by the volume standard deviation.

B.3.8 Volume differential peak density

As Volume Peak Density but taken on the first differential of the volume.

B.3.9 Volume trough length

The average durations for which the volume is lower than half a standard deviation below the volume mean.

B.3.10 Volume trough minima

The mean of the volume minima of volume troughs divided by the volume standard deviation.

B.3.11 Dynamic profile

In addition, the profile of expenditure of energy (precipitous for high arousal, smooth for low) before and in between onsets, which appears to be important mirror neuron information, will in future be predicted by computation of profiles of energy flow leading to significant articulations.

For example, τ"tau" coupling (Lee 2005): $\tau_x = K_{x,g} \tau_g$

where tau = time at origin of glide (end of previous onset), x = the gap preceding the next detectable onset, g = a patterned flow of electrical energy through an assembly of neurons, kappa = movement value determined by the brain. Profiles of energy will be determined by profiles of mean values of kappaXG.

B.3.12 Standard, commercially available software for rhythm detection may be used satisfactorily for some genres of music, but such software may fail to detect the specific bio-activating rhythm of any given piece of music and may even have difficulty in detecting rhythm at all in some. The above algorithms, which predictively model the activations of core rhythmic processing centres of the brain, have proved reliable. Some of these algorithms, for example beat detection, could in theory be replaced by other mathematical procedures. The originality of the invention lies in the unprecedented nature of the biological modelling. Thus we have a phenomenon in music (rhythm) that is known to have an effect on arousal and counter-arousal in the autonomic nervous system (as well as core emotional systems, endocrine activity and neurotransmission), which in turn is known to have a powerful influence on how you feel: relaxed, able to concentrate, wanting to dance etc. We also have a means of measuring the effect of the rhythm (our sensor). Our categorisation algorithms (above) take as an input the relevant data from the digital signature analysis and yield as an output a predicted impact on the chosen biometrics. Intense rhythms will have an arousing effect while gentle rhythms will have a calming effect, and there is no shortage of prior art based on the same principle. In modelling the innate neurophysiological response to rhythm an algorithm linking this measurement of rhythm to its expected effect on (in this embodiment) heart rate and galvanic skin conductance is hypothesised, tested and refined.

B.4 Turbulence and core emotional systems (Locations and Organs)

The 'turbulence' of a piece of music relates to the speed and extent to which it changes over a period of time, in terms of rhythmicity and harmonicity as well in terms of general fluctuations in sound pressure.

'Turbulence' combines indices of change in rhythmicity and harmonicity, related to pathways described above, with auditory brainstem and cortical activity innervating the amygdala, hippocampus and core emotional regions affecting neurotransmission and endocrine systems, including the HPA axis, dopamine circuits and levels of, for example, norepinephrine, melatonin and oxytocin (Miluk-Kolasa et al 1995; Gerra et al 1998; Kumar et al 1999; Evers and Suhr 2000; Schneider et al 2001; Blood and Zatorre 2001; Grape et al 2003; Uedo et al 2004; Stefano et al 2004; Herbert et al 2005; Nilsson et al 2005). This important predictor of arousal and counterarousal may be represented as the differential of rhythmicity and harmonicity.

'Turbulence' is therefore a measure of rate of change and extent of change in musical experience. These factors seem to activate core emotional systems of the brain, such as the amygdala and periaqueductal grey, which are in turn linked to autonomic and endocrine systems. At high levels of musical energy turbulence may enhance arousal; at low levels it may add to the counterarousal effect.

The total turbulence (T) of a piece is determined as a combination of the turbulence of the harmonicity (H') of the piece and the energy present during peaks of volume of the track (P). Turbulence of harmonicity is calculated as the standard deviation of the differential of the harmonicity, divided by the mean of the differential.

In one very simple implementation, total turbulence is calculated as:

T=dH/dt * P

This equation is a non-limiting example of how turbulence can be calculated and other ways of linking T to H and P may well be appropriate; furthermore, T may be defined in terms of other or additional variables.

See **Figures 14** and **15**, showing volume and harmonic energy as a function of time.

B.5 Combining values

Each of the algorithms described above, hypothesised and refined through testing, has effectively become a 'virtual organ' of the brain that helps us predict the effect on levels of arousal and counter-arousal of patterns that can be detected in music using digital signature analysis. The relative weighting of each 'organ' may be adapted using heuristic, machine learning or other techniques to calibrate the overall predictive power of the set of 'virtual organs' working in harmony.

Any subset of the above analyses may be combined together to produce a single number estimating where a piece of music (or part thereof) lies on the scale from relaxing to exciting. The formula used to perform this combination may be derived from experimental data, as follows: A number of human listeners listen to the same selection of tracks. Each listener then independently ranks all the tracks in order from what they consider the most relaxing to the most exciting. (The ranking could also be done objectively by measuring the listeners' physiological data, but this has so far given much less consistent results across listeners.) A statistical regression analysis is then carried out, with the average human ranking as the dependent variable, and the chosen subset of musical analyses as the independent variables. In other words, a single formula is produced which uses the analyses to predict the human rankings. The coefficients in this formula are chosen to give the best possible prediction, considered over all tracks. The resulting formula may then be used to produce automated predictions on a mass scale for a much larger number of tracks. Consider the following example data:

Track	**Average human ranking (0-1)**	**Mean harmonicity (mh)**	**Volume level (vol)**	**Rhythmicity (rhy)**
1	0.2	0.212	0.010	118
2	0.4	0.231	0.069	228
3	0.5	0.204	0.123	187
4	0.6	0.225	0.294	130
5	0.8	0.173	0.163	155

Any statistical regression method may be used to produce the overall formula. For example, if we use multiple linear regression with the ordinary least squares estimator, we obtain the following:

$$\text{Predicted ranking} = -6.59*\text{mh} + 1.63*\text{vol} + 0.0018*\text{rhy} + 1.36$$

Non-linear transformations of one variable (e.g. logarithm or reciprocal) or non-linear combinations of multiple variables (e.g. their product or ratio) may also be used, by pre-calculating them and then treating them as additional variables in the regression.

The coefficients employed in each of the algorithms, and the relative weighting of the algorithms in combination, may be optimised for different musical styles using metadata (such as genre and artist) that are typically carried alongside music distributed in digitised formats such as the Compact Disc and over the Internet. With the accumulation of large amounts of (anonymised) human response data that may be fed back (with the consent of the listener) in networked deployments of X-System it will be possible to fine-tune the relative weighting of both the equation coefficients and their relative weighting in combination to improve accuracy. Similar optimisation of coefficients and weightings will be achieved by analysing user data in combination with the music metadata (such as genre and artist) that are typically available with music distributed in digital formats, and in due course this optimisation will be extended to both the individual user and specific recordings.

The overall arousal index calculated for each piece of music may be expressed either as a single number that describes the overall neurophysiological effect of listening to it from start to finish, or it can be displayed graphically with arousal index on the vertical axis and time on the horizontal axis. The resulting trace would effectively describe the neurophysiological journey a listener may expect as they listen from beginning to end. This latter is likely to be of particular use in longer and more complex pieces of music such as much of the classical repertoire, whereas some other repertoire such as modern Western pop music might more conveniently be represented by a single number. In either case, the effect of a piece of music is both inherent (in that it is a product of the patterns detected in the music) and dependent on the state of the listener (in that the neurophysiological effect of music is relative rather than absolute [Altshuler 'The Iso-Moodic Principle' 1948]).

As we learn to navigate the brain in greater depth and detail, and as sensor technology develops further, different equations will be developed to predict the effect of different musical structures on different measurable outputs. All these instances of the application of the Innate Neurophysiological Response to Music are intended as different implementations of the present invention, which claims a novel system and method of predicting the effect on universal human neuro-physiology of any piece of music from any musical tradition by means of analysing bio-activating patterns in music and using mathematical equations tailored to specific biometric indices to predict the effect of these musical patterns on the chosen biometric indices.

B.6 This section describes an alternative approach to combining values for rhythmicity, inharmonicity and turbulence to produce an excitement (E). In this alternative approach, E is given by:

E=(10*I*R)+T

This equation is a non-limiting example of how excitement E can be calculated and other ways of linking E to I, R and T may well be appropriate; furthermore, E may be defined in terms of other or additional variables.

This generally produces a number from between -1 and 7, representing the range of the counterarousal-arousal scale. Currently the thresholds for five arousal categories are approximated as

-1 to 0.6 = 1
0.6 to 2.2 = 2
2.2 to 3.8 = 3
3.8 to 5.4 = 4
5.4 to 7 = 5

An alternative is an equation where rhythmicity and harmonicity are multiplied and turbulence added. In other examples, log scales and Fibonacci progressions may be used in the analysis of auditory data.

More detail: For each of R, H and T, X-System records both a single average value (μR, μH, μT) and a profile of variation further categorized as ascending, descending or stable ($\Delta R>0$, $\Delta R<0$, $\Delta R=0$; $\Delta H>0$, $\Delta H<0$, $\Delta H=0$; $\Delta T>0$, $\Delta T<0$, $\Delta T=0$).

The average values of R, H and T are mapped (in the simplest case the normalised mean is taken) to an n dimensional point p characterising physiological state. The variations of R, H and T are also mapped (again, in the simplest case the normalised mean is taken) to another n dimensional point q characterising the directional effect these values will have on the physiological state.

The concatenation of p and q allows each musical excerpt to be mapped onto a Musical Effect Matrix M, a 2*n dimensional matrix, n dimensions corresponding to the physiological parameters measured by E representing granular ranges into which E can fall, the other n dimensions corresponding to the effect the track will have on the physiological parameters (ascending, descending or maintaining any given physiological parameter or dimension of E).

We now describe in more detail how the Music Effect Matrix M is generated. As noted earlier, **Figure 7A** is a block diagram showing the major components in X-System for analysing harmonicity and **Figure 7B** is a block diagram representation of all of the major components of the musical analysis tool. The values output by the analysis are specified as functions in t, the time index of a particular measurement. These values (corresponding to R, H and T) are grouped as follows:

X(t): values for rhythmic "presence", tempo, power and density of pulse-related rhythmic structures, and harmonic rhythm - related to cerebral cortex activity, core emotional locations, and autonomic and endocrine responses.

Y(t): degree of conformity, within the limits of human perception, to exponential series-related frequency structures in melody and harmony - related to the cochlea, Heschl's gyrus and cerebral cortex processing, core emotional locations and autonomic and endocrine responses.

Z(t): the rate and magnitude of variation in X(t), Y(t) and dynamic power (W(t)) which is measured using the normalized, gain adjusted volume level - related to activation of core emotional systems, and the endocrine and autonomic nervous systems.

Categorization may be preceded by aggregation, documenting provenance, genre and other data for music tracks. This may be according to an industry standard such as that provided by Gracenote®, it may be the result of individual user editorial, crowd-sourcing methods such as collaborative filtering, or may be the result of future aggregation standards based on, for example, digital signature analysis. The purpose of aggregation is to allow the user to choose a preferred musical style, though it is not strictly necessary for the proper functioning of X-System.

In order to reduce the computational cost of analysing a piece of music, only certain regions are examined. The location and length of these regions are determined dynamically, based on configurable parameters and an adaptive mechanism that recursively examines regions with a large rate of change. This produces a sparse array of values for each function, identified by a time index. Due to the recursive analysis, the step size _t will vary over the function domain t.

Algorithmically, these regions are generated by applying a windowing function to the incoming audio data. The sampling window is then “stepped" over the region, and the results of each step are aggregated to form the single output at time t. For example, a region may consist of the (absolute) time interval (0s; 1s), which is further windowed into 50ms samples, with a 10ms step size. This produces a total of 96 sample points, which are combined to form a single value X(0) = x.

The analysis of X(t) is performed by an "acoustic streaming"- based rhythmic induction, combined with pattern-recognition and an index of power and density.

Rhythmic induction is performed using two main techniques; band-limited power spectral density onset analysis, and adaptive comb filtering. The results of both techniques are then subjected to a number of heuristics based on music theory, and are combined to form a single estimate of the musical rhythm.

Heuristics include rules such as the minimum and maximum plausible tempos or some form of probability distribution of likely tempos for a given input genre if known. They may also include emphasis and de-emphasis of certain frequency bands based on the input.

Spectral Density Onset Analysis uses a sequence of short-time Fourier transforms of the windowed samples to calculate the energy present in specific frequency bands. This data is tracked temporally to observe peaks in bands, which characterise rhythmic events.

Comb Filtering involves convolution of the input signal with a variety of impulse trains of different spacing, on the basis that as the impulse spacing approximates the rhythm of the input, the overall convolution result will increase. This technique can then be used recursively to _find a best-fit impulse spacing which characterises the input rhythm.

Values for Y(t) are established by means of an adaptation of auditory scene analysis. The audio input data are passed through a gammatone cochlear filter bank, splitting them into multiple streams. For each stream, special, frequency and onset information is calculated.

Spatial information is acquired from stereo tracks of each stream, frequency peaks are calculated using a Fourier transform and onset detector maps are applied to find the starts of sound elements.

This information is combined and correlated to partition the audio data input into sound sources. For each of these sound sources a number is calculated as the ratio of sound energy within the harmonics of its fundamental frequency to the sound energy outside the harmonics of its fundamental frequency. Y(t) is the mean value of the ratios for each sound source from the excerpt.

The fundamental frequency is determined using a Harmonic Product Spectrum, in which the signal is repeatedly multiplied with down-sampled copies of itself, causing a large peak to occur in the frequency spectrum corresponding to the fundamental frequency. Standard signal-processing techniques are also applied to de-noise the resultant output.

$Z(t)$ is measured as the rate and magnitude of variation in $X(t)$, $Y(t)$ and $W(t)$.

In each of these cases ($X(t)$, $Y(t)$ and $Z(t)$) the system records both a single average value (μX, μY, μZ) and a profile of variation further categorized as ascending, descending or stable:

- Ascending - An overall positive trend in the functions $X(t)$, $Y(t)$ and $Z(t)$.
- Descending - An overall negative trend in the functions $X(t)$, $Y(t)$ and $Z(t)$.
- Stable - Only minor deviations from the mean μ result over the audio input signal.

The average values of X, Y and Z are mapped (in the simplest case the normalized mean is taken) to an n dimensional point p characterizing physiological state. The variations of X, Y and Z are also mapped (again, in the simplest case the normalized mean is taken) to another n dimensional point q characterizing the directional effect these values will have on the physiological state.

The concatenation of p and q allows each musical excerpt to be mapped onto the Musical Effect Matrix M, a 2n-dimensional matrix, n dimensions corresponding to the physiological parameters measured by E representing granular ranges into which E can fall, the other n dimensions corresponding to the effect the track will have on the physiological parameters (ascending, descending or maintaining any given physiological parameter or dimension of E).

C. How X-System is used

As noted above, X-System may use a subject's biometric data (where a sensor is available) to measure neuro-physiological arousal. It then leads the subject by stages towards a target level of such arousal, state of mind and/or affect. This is achieved with a database of music, previously categorised using predictive modelling of innate neuro-physiological responses. Categorisation in real-time or near real-time is also possible. Categorisation can be visually displayed (e.g. on the display of the computing device used for music playback); this can include a display of the E values for each music track, or how the E

(Excitement) value changes during a track; R, I, H, C and T parameters can also be visually displayed. A piece of music that predicts or matches the subject's current level of neuro-physiological arousal is selected and a playlist constructed on the basis of the fundamental musical effect of each constituent piece of music. Listening to the playlist directs or moves the user towards the desired level of arousal, state of mind and/or affect by unconscious neuro-physiological entrainment with the music and enables that level to be maintained. The subject's current level of neuro-physiological arousal can also be visually represented, as can the convergence to the desired target state.

X-System is, in one implementation, designed to sense the state of mind and body of the user and stream music of selected repertoires to achieve target states such as:

• Excitement
• Relaxation
• Concentration
• Alertness
• Potentiation of physical activity

See **Figures 2, 3 and 9**, for example.

C.1 Components in the X-System

X-System includes:

- automatic categorisation software capable of categorising music of all cultures either remotely or in proximity according to specific levels of arousal and counterarousal; these categorisations may be offered for general use independently of the sensors and diagnostic software. This may be based on Nigel Osborne's INRM (Innate Neuro-physiological Response to Music) paradigm.
- a database of music categorised manually or automatically (using the automatic categorisation software) to achieve specific levels of arousal and counterarousal
- sensors to detect physiological indicators of arousal (such as excitement) and counterarousal (such as drowsiness), including heart rate and galvanic skin conductance

- diagnostic software which employs sensor data to monitor levels of arousal and counterarousal in the user
- music playback/streaming (eg. playlist selection) software which selects previously categorised music from a database to stream appropriate repertoire to achieve target states of mind and body by a process of step-by-step entrainment, starting from the current diagnosed "state"; progress towards these goals is monitored by the diagnostic software. Specific tracks for a listener may be selected for playback (by streaming or otherwise) according to bio-feedback from that listener; the playlist may be created locally and the music tracks requested for streaming/download etc; it is possible also for the bio-feedback and desired "state" information to be sent to a remote music server and for that server to generate the appropriate playlist and provide music tracks to the local, personal playback device. In this variant, the personal playback device need have no local music library or X-System software/firmware etc.,; it needs only the ability to detect the listener's audio preferences and bio-feedback data and to relay that back to the remote server using a low capacity back-channel and to then receive the music from the remote music server.

Note that all software may also be implemented in hardware, firmware, SoC, as part of a third party audio stack and in any other convenient manner.

Appendix 1 is a more detailed description of the components of X-System.

C.2 Practical applications of X-System

The sensor is intended to measure one or more predetermined parameters of the user's state of mind and body and to communicate this information to a processor; the processor is designed to select tracks from the music categorisation data appropriate to lead the user from her/his current state of mind and body to the intended condition of arousal or counter-arousal. This combination will allow X-System to:

- Sense, in real time, the neuro-physiological state of the human mind and body;

- Analyse the music collection of the consumer;, or any other collection he/she has access to, such as with a cloud-based or remote/central server based music service; and
- Calculate and deliver play lists as a function of a desired state of arousal.

This will enable users to direct themselves to a desired state, such as:

- Excited and ready to play sports or exercise; for example, to enhance oxygenation levels for competition or reduce post-surgical recovery times;
- Relaxed and able to drift off to sleep;
- In a meditative state to support development of insight;
- In a meditative state to support the development of creative thought; and
- Maintaining focus and able to concentrate.

(for example, to provide support to overcome conditions such as insomnia, to reduce medication in post-traumatic stress disorder (PTSD) and in mania patients, to develop and to organise memory, categorised by short, medium and long term need for data retention.), and to create a state in which to encourage creativity and imagination.

The diagram of **Figure 4** illustrates the current project implementation of X-System. In an alternative to the implementation of **Figure 4**, because ubiquitous mobile computing blurs the distinction between devices, the elements shown in **Figure 4** within a User PC (music player, music library, automated analysis and musical effect database) may be distributed over two or more computing devices. In a commercial example it may also be configured to work with portable audio devices: see **Figure 5**.

While these components are key elements of X-System, its core innovative technology is a definition of the bio-active components of music (based on a predictive Innate Neuro-physiological Response to Music paradigm, Osborne 2009, eg. see **Figure 1**), the algorithms used to calculate them based on digital signature analysis and the calibration methods used to tune the system to the neuro-physiological response of an individual.

D. The Sensor or Sensors

The sensor may be in the form of a wristband, a hand-held or any other device suitable for taking the required parameter measurements. The sensor may be body-mounted, or use ear buds (e.g. combining a sensor into ear-bud headphones), remote monitoring via IR or acoustic, wireless, or more generally any form of life sensing. The data captured preferably comprises biometric parameters such as heart rate (including pulse rhythm analysis), blood pressure, adrenaline and oxytocin levels, muscular tension, brain waves and galvanic skin conductivity. Alternative equipment formats include necklaces, bracelets, sensors embedded in clothing, other jewellery, sensors implanted under skin, headsets, earphones, sensors in handheld form such as covers for 'phones, MP3 players, or other mobile computing devices.

Sensors currently used in the X-System project comprise a wristband sensor which will be used to measure galvanic skin response (GSR), and a standard finger clip Pulse Oximeter for the measurement of heart-rate and blood oxygenation. For the purposes of commercialisation these sensors will be combined in a single, wearable, wireless device. Other potential bio-sensors and motion sensors may be included as they become economically viable.

The sensors must be able to measure a combination of pulse rate and skin conductivity, combined with any other possible measurements and must be resistant to disruption from movements of the user or changes in environment; it must also be possible to wear the sensor for extended periods of time without discomfort or embarrassment. Other sensors include physical bio-sensors such as oxygenation, EDA, EDC, EDR, ECG, sugar levels, BPM, EEG etc, and multi-spectrum sensors (radio, IR, UV, heat, and broad spectrum), which detect bodily radiation auras.

Figure 5 shows a desired architecture overview. Figure 5 shows an implementation of the X-System invention where a primary music library, and analysis software resides on a user PC that is operable, remotely or locally by the listener or a third party, with the ability to transfer a selection of music to a personal music player device, which then generates a dynamic playlist based on the available music.

The X-System sensor measures certain chosen parameters of the user's physiological state and transmits the resulting data wirelessly to a processor in (or in communication with) a playlist calculator, which resides on or is otherwise connected to a music playback device (for example, a personal computer, smartphone, MP3 player or other audio device). Transmission is preferably wireless but it will be appreciated that other transmission types are possible. Indeed, the processor may be integrated with the sensor.

The chosen physiological state parameters are denoted by P. A function F(P) reduces these parameters to a single, normalised point E, characterising the general physiological state of the user. In the simplest case E is a one-dimensional measurement of the user's physiological arousal (or counter-arousal). With further inputs a more complex measurement may be obtained, resulting in a point E of n dimensions. An effective prototype has been developed using pulse rate 'p' and galvanic skin conductivity 'v' to calculate a simple index of physiological arousal where $E = p+v$. Currently the prototypes use the Nonin X Pod Pulse Oximeter and a skin conductance biosensor. The pulse rate, oxygenation and skin conductance of the user are constantly monitored; heart rate may be used as to control mean variations in conductance. Both sensors currently work independently and are connected wirelessly to a controlling computer. They may be replaced with a single integrated sensor. Alternatively, any other form of wired or wireless communication of sensor outputs to player to output device is possible. **Appendix 1** gives more details.

A user initially provides the system with their personal music collection (or uses an online library of streamable or downloadable music). This is analysed for level of excitement, using INRM categorisation in combination with signal processing and machine learning techniques. The user then synchronises this information with their music player and selects a level of excitement/arousal; someone other than the user may also select the excitement level. The sensor wristband provides the system with a constantly updating real-time state of excitement of the user, allowing the system to react to external effects on the user and "catch" them, using the principles of entrainment to bring them back towards the desired state. Once the user has achieved the target level of excitement, they are kept there by music determined to be effective at maintaining that state.

Although the current version of X-System's sensor is based on heart rate and skin conductance, there are strong arguments for early integration of other measures, including for example EEG, brainwave sensors. This would allow factors such as concentration, alertness, contemplation, drowsiness or creative flow to be monitored directly through sensing of frequencies of entrained firing of neurons in the brain, rather than indirectly through indicators of arousal. A second set of related challenges lies in further aspects of machine learning. Individual physiological responses vary considerably, from person to person, according to time of day, state of metabolism etc. X-System may learn from individual users the range of their physiological responses in order to identify relative levels of arousal, and individually calibrate the diagnostic software. It may also learn about their personal preferences as already articulated through their choice of repertoire. X-System may also go directly from a set of musical features, using a neural network to predict the effect of these on physiological measurements, without first reducing the features to an expected excitement/arousal level.

E. Musical Selection Algorithms

Certain levels of neuro-physiological arousal are necessary precursors of activities such as sleep, relaxation, accelerated learning and study, or increased alertness and activity. The user will preferably be presented with a user interface and choose from a menu of such activities in order for the system to establish a target level of arousal and affect that will facilitate the chosen activity.

The point E, representing the neuro-physiological state of the subject diagnosed by the sensor, is used to select music from a database of music tracks indexed by the Musical Effect Matrix M, based on a combination of the granular point r and a direction d pointing towards the physiological state towards which the user has elected to move (see preceding Section E for more detail).

The first piece of music selected will correspond to the initial neuro-physiological state of the subject, represented by E. Subsequent pieces are selected based on their values in M such that each would, played in order, be capable of progressively leading the subject's state towards the target state. The order in which the pieces of music are eligible to be included in a playlist is determined by a vector that represents a temporally-organised ascending, or descending as appropriate, series of musical effect values in M. The set of pieces of music in the database that meet the requirements of this series of effect values is known as 'Qualified Content'.

The Qualified Content is arranged into an actual playlist according to a set of rules, including but not limited to random selection, anti-repetition, genre preference or some other heuristic. In some cases it may be appropriate to comply with the US Digital Millennium Copyright Act (DMCA).

Where a sensor is used, then a biofeedback loop is established in order to ensure continual recalculation of the playlist to compensate for distraction, individual sensitivity and other factors based upon any dimensions of overall affect that are susceptible to continual measurement. Direction towards non-measured parameters of state of mind and/or affect will still occur despite the lack of a bio-feedback loop because neuro-physiological arousal is a necessary precursor to state of mind and affect and establishes the conditions under which the listener is most susceptible to these other aspects of overall musical effect.

Once a piece of music has been played it is preferably removed from the list of potentially available content for a minimum number of cycles in order to avoid unnecessary repetition. This anti-repetition rule is subject to a feasibility test in order that a message of appropriate severity may be displayed to the user warning of insufficient content or variety of content in the music database to enable effective functioning of the system along with a suggested remedy such as a recommendation of further pieces of music which might be added to the database to improve its functioning.

In the case where content has been distributed pre-categorised or where it is streamed from a central server, playlists may be calculated initially in a dynamic mode where shorter excerpts are taken from the database. Once the listener has achieved the target level of arousal, longer excerpts are admitted into the qualified content pool for the purpose of playlist calculation and the system may enter maintenance mode. Any disturbance which causes the listener's level of arousal to vary by more than a predetermined factor may cause the system to re-enter dynamic mode and re-calculate the playlist based upon shorter excerpts in order to entrain the listener back to the target condition at an accelerated rate.

The anti-repetition rule as applied to shorter excerpts may be used to calculate the minimum required catalogue size on the basis of the number of separate musical styles that may be selected by the user, the average length of a shorter excerpt, the minimum number of cycles that must pass before the anti-repetition rule will admit a song or excerpt back into the selection pool and the number of shorter excerpts available that fall within the least-populated cell of the musical effect matrix.

F. The Music Player

The music player may be an adaptation of standard industry software such as the Windows Media Player which is capable of building dynamic playlists according to the Musical Selection Algorithms and of offering the user additional utility such as selection of musical style, display of associated metadata and video content.

The music player may also be a software application which is downloadable from a software application store accessible via the internet. **Figure 8** summarises a design of the player system and the integration with the sensor subsystem. In an implementation, a player system and subsystem may be distributed across two or more computing devices; ubiquitous computing methods allied to mobile computing and personal human inputs may be employed, together with multiple ways of processing and delivering audio outputs, both private and public. So not only players, but also processors and human interaction devices, including but not limited to entrainment of interaction and control of a personal environment by emotional cues, as well as ordering or sequencing consumption may be used in an implementation.

G. Diagnostic and streaming software

When a sensor is used in System-X, then diagnostic and streaming software is capable of reading the values from the sensor(s) and determining a state of arousal of the user. The nature of skin conductance means that the absolute value can vary significantly due to how well it is in contact with the skin, from person to person and through normal sweating. To rectify this, the skin conductance value may be calibrated automatically based on the heart rate of the user.

The user of the system wears the system, selects a repertoire of music that they would like to listen to, decides what excitement level they would like to get to and puts on the sensors. Once a diagnosis has been made for the state of arousal of the user, this data along with the selected excitement level is used to select a program of tracks from the repertoire.

Optionally, the user selects a repertoire of music e.g. Jazz, Classical, Indian, World, Baroque), decides what their target arousal level should be (e.g. relaxed, excited, steady) and puts on the sensors. Once a diagnosis has been made of the current state of arousal of the user, repertoire is automatically selected to lead or "entrain" the listener from their current state to their chosen state of arousal. This is performed by defining a playlist, which entrains the user from the current emotional position in the multi-dimensional space defined by the INRM parameters, moving in small steps towards the defined position in INRM space defined as the desired end point.

H. Manual categorisation

In an example, the repertoire has been categorised manually by a combination of pulse/metre detection using a metronome, and intuitive predictive judgements concerning levels of arousal and counterarousal associated with various musical parameters including rhythmicity, harmonicity, turbulence etc. e.g., the faster the pulse/metre the higher the arousal, the higher the harmonicity the lower the arousal. In the sample categorisation of **Figure 16** (from the Miles Davis repertoire) tracks are placed in one of five categories corresponding to levels of activation/arousal.

I. Manual categorisation vectors

By way of example, in other manual categorisations tracks are further sorted into stable, rising and falling vectors, e.g. "category 4 rising" will be selected if the user chooses a target state of high activation/arousal; "category 4 stable "would be selected if the uses wishes to remain in a state of moderate activation. For an example, see **Figure 17**.

In the example of **Fig. 18**, movements from Beethoven symphonies have been categorized according to the vectors. Note that no movement was identified as appropriate for 4/stable or 2/stable.

Examples of the present invention have been described with reference to its effect upon human beings. However, the effect of music on animals is well documented. This almost certainly depends on simple psychoacoustic effects of sound environment, rather than a musical/biological discourse as such, but examples of the present invention may see applications in animal husbandry or veterinary medicine in addition to both general consumer, professional, athletic, wellness, healthcare and other markets.

J. Social Networks

In this application, X-System is adapted to facilitate the communication of neurophysiological state, arousal, affect and valency data, determined by X-System's algorithms, to friends via short range wireless and Bluetooth networks, as well as more widely to social networks such as Facebook and Twitter, and to health and care workers, as a diagnostic, monitoring or entrainment tool.

This application enables a range of navigational and communication applications on smartphones and other devices, allowing users to 'communicate and navigate innate states of arousal' (mood or emotion) and 'communicate and navigate experience'. It enables individual X-System users not only to see displays showing their own innate states, but to allow others to 'read' their true or unconscious states as they experience a variety of activities, from listening to music, to sports and recuperation and post-surgical care in health care settings.

A system and method for communicating X-System diagnostic capacity to decode neurophysiological states, adapting it to facilitate deeper, more direct communication about states of arousal and valency whilst engaging in a wide range of activities (including but not limited to music), between individuals and groups in social networks.

A system and method for generating information requests based on actual states of arousal (as measured by X-System), to search engines such as Google – this arousal information can then be used as an input to the search algorithm and also to the algorithms that control which advertisements are displayed (so for example, web users may be more receptive to advertisements for certain products when in specific arousal states and search results and advertisements can be tailored for maximum relevance using arousal state information. The arousal information can be used also to indicate 'presence' status information ("I am in a good mood, listening to Beethoven" etc.).

X-System categorises the innate neurophysiological 'state' of individuals in terms of both an unbroken continuum of data and discreet categories, ranging from 1 (high arousal) to 5 (counter-arousal). This is linked in core X-System applications to music selection.

In this 'social networking' or 'sharing' application, the innate 'state' arousal/counter arousal and valency data of an individual is transmitted over a variety of standard communication networks (including but not limited to Wi-Fi, Bluetooth, GSM, and other Mobile networks and fixed-line Internet) both directly and via wider social network systems (such as Facebook), to enable peer to peer and one to many communication of arousal, together (optionally) with coding that indicates concurrent music or other entertainment selection, or self-declared activity ('this is me watching a movie; responding to an advertisement; walking in the country; running, cycling'), all in real time, or near real time. For example, X-System detects emotional arousal parameters information of an audio track and then embeds this information into the audio track or into an electronic link to the audio track or as metadata associated with the track.

The X-System 'state' data can be distributed in real time snapshots (arousal and valency now); in real time streams (continuous flow); as history (arousal and valency yesterday), with or without data about the music selected at the time. This might be termed "a personal verve index" (verve: vivaciousness; liveliness).

The data will then be displayed as graphics, as colour codes, or in a variety of statistical forms. Users will be able to annotate the data and music with 'activity labels' (I was running at the time, or doing homework), which will open up other forms of analysis of the relationships between arousal, valency, music, other entertainment experiences and activity.

This application will enable individuals to search for people in their social networks who are in a similar mood, or engaged in similar activities, such as 'find people in my network who want to talk' or feeling down and yet keen to talk. This can be indicated by mood boards or to augment presence information on Facebook and other social networks.

With large volumes of users expressing their mood automatically generated by people who opt in (subject to say anonymity rules and permissioning about sharing), the data can indicate overall states of arousal amongst groups and larger communities.

The application will be extended to provide graphical and network maps showing patterns and cluster of moods amongst social groups, creating a 'social emotion' landscape for groups either engaged in their own individual activities, or groups together, in a social setting, such as at a party, or listening to a concert, or dancing.

This contrasts with early examples of social network analysis, which are limited by data mining and pattern matching derived from language and semantic analysis and so limited in their accuracy. X-System will generate more authentic and accurate interpretations of both individual and group arousal by capturing true innate neurophysiological state information.

This application will also be used to optimise web sites by linking X-System users to web cookies, such that if I am looking at a site and agree to the use of X-System readings of my innate state information, the cookies will generate analysis of the emotional impact of the site, or particular pages. This will enable web designers to experiment with a variety of textual, film, music and screen displays, layouts and experiences and get immediate feedback about users' emotional response.

This information will then be available to be matched to advertising and marketing metrics, such that responses to web experiences can be aligned with brand values and with the desired moods or desires that particular products and services aim to create. So, for example, the feedback mechanism might be used to match the emotional response to an advertisement about a particular car.

This extension of X-System's core algorithms creates a new form of communication, operating at a deep level, beyond culturally bound, linguistic expressions of mood, optionally linking it to current activity including choices of music, other entertainment and other activities.

This communication of unconscious, pre-linguistic levels of arousal, affect and valency opens up a new paradigm for social networking and health care diagnostics. In care settings, for example, monitoring patients' 'state' information will provide insights otherwise not possible using conventional diagnostic techniques. X-System may be integrated with a variety of conventional medical, care and diagnostic devices and applications to create a more holistic picture of of patient condition and emotional state.

The X-System core data about innate arousal, valency and music selection is transmitted via standard interfaces to widely available social networks such as Facebook and Twitter and direct to Smartphones in local networks.

X-System will be embedded in Smartphones and other devices, in a variety of combinations of software, firmware and chip hardware. The X-System API will enable specialist App developers to create a variety of tools and techniques to leverage the flow of 'state' information, creating feedback and monitoring services.

There are many protocols and systems for the transmission of data and interfaces to social networks and Smartphones. This application of X-System is unique in that it enables these systems to be extended with new data that is otherwise not available. X-System is extended to target communication of innate arousal and valency with accompanying data indicating concurrent music, other entertainment or self-declared activity to individuals and groups in local, wide area and social networks.

X-System can also share arousal values, associated with a user interacting with a search engine such as Google®, with that search engine. The search engine can then use those values to optimise search and/or advertisement selection by the search engine.

X-System can also share arousal values associated with a user browsing a specific website or pages in a website with a website optimisation system so that the website optimisation system can use those values to optimise the website and/or specific pages (content, layout, sounds etc.).

K. Opportunities for Expansion/Enhancement

The main directions of product improvement and expansion are as follows:

- Identification of emotional responses to music stimulated by memories or response to lyrics or other aspects of a song or piece of music rather than biology - developed by filtering out the expected physiological responses.

- Sensor development and accessories, such as new generations of miniature Electroencephalography (EEG) brain scanning sensors. One possible approach is to include sensors (measuring any of the parameters discussed above, such as pulse, skin conductance etc) in earbuds or in gloves.

- Advanced music search, navigation and discovery systems.

- Advanced music search, navigation and discovery systems, including promotion, ordering, selection, and control interfaces.

- Specialist medical applications.

- Analysis of music to determine innate emotional responses; and

- Capture and analysis of sensor data from early adopters to fine-tune level of arousal.

There are two further strategies for refining analytical functions. The first is through large-scale usage of the system. It is proposed to recruit one hundred volunteers to test the system in five phases. Their physiological data during listening, including heart rate and skin conductance readings, will be compared with automatic categorisation data and the results of manual categorisation, as a means of identifying strengths and weaknesses in the automatic analysis process, both in the capture of data and in the combination of values.
The second strategy for refinement is through machine learning, making use of linear regressive and/or neural network approaches. Training phases will follow each of the five testing phases. This approach will have the value of both scrutinising existing values and their combinations, and building up an evolving resource of learnt information and

procedure. It may not be possible to refine the automated classification significantly. If this proves to be the case, machine learning processes and statistical analysis will be used to generate the necessary refinement. Additionally, weaknesses in the automatic classification system can be corrected through gathering and analysing the actual measurements of the effects of specific tracks on users. Those skilled in the art will appreciate that both artificial intelligence (AI) and heuristic rules-based approaches, and iterative automation and testing methodologies, may be employed.

X-system could also be used to create and adjust 'mood' in retail environments, and/or in online communities, through the playback of suitable music. Individuals could be connected via web interfaces to generate a common response/reading.

Similarly, X-System could be used in the understanding of and the matching of emotional responses to brands – essentially using X-System as a tool by which to diagnose and then shape emotional responses to brands by associating those brands with exactly the right kind of music for the target audience. X-System can be used in judging the response of different social groups to brand music.

Using polling or similar crowd-sensing techniques, X-System can also be used as a dynamic group entrainment tool in group environments, to select music which heightens arousal, for example at sports or entertainment events, and to reduce group tension and frustration in public environments such as transport, hospitals and government buildings.

L. Benefits of X-System

This technology is anticipated to have broad social, psychological and biological benefits in the reduction of stress, the treatment of insomnia, in optimising concentration and learning, in improving creative thought, and in facilitating optimal exercise patterns, whether for the general population or to support training regimes for elite athletes, and enhance event competitiveness.

X-System may be applied in therapeutic approaches to specific medical conditions. There is a large body of literature that provides evidence of the efficacy of music medicine and music therapy as complementary support in the treatment of conditions such as chronic pain, dementia, Parkinsons disease, depression, post-traumatic stress disorder and aphasia, and in palliative, post-surgical, post-stroke care. Possible benefits include reduction of bed rest after surgery, and reduction of drug use.

As an example, Jane would like to be able to concentrate better on the task at hand, so she slips on the wireless sensor wristband, touches the "concentrate" symbol on her iPhone and listens as she gets on with her work. The system will monitor her state of mind and body and play music suitable for maintaining an appropriate level of concentration.

It should be noted that in addition the automatic categorisation algorithms of X-System have considerable potential market value as a "stand alone", independent of the sensor technology, capable of offering an "emotional" navigation capacity for music streaming systems.

The invention may be used beneficially to select and categorise music according to its neuro-physiological effect, including but not limited to the ordering/sequencing, use, promotion, purchase and sale of music according to its neuro-physiological impact. The invention may also be used beneficially to link such categorisation to other categorisation schemes in common use.

Other potential uses of this system could be for selecting appropriate pieces of music from a database of library music for the soundtrack in films where a specific mood of the viewer is desired. It could also be used in visual arts, where a specific mood of the viewer is desired. Hence these applications would be visual applications or audiovisual applications, rather than just audio applications.

Related products and services will be generated from both of these areas to generate market intelligence about future trends in markets, i.e. products and services relating to analysis of music to determine innate emotional response, and capture and analysis of sensor data from early adopters to fine-tune level of arousal will be generated to generate intelligence about trends in future markets. Examples may include services to the computer game industry to assist in sound track selection to enhance the emotional experience of interactive gaming technology or as an aid to music composers seeking to elicit a particular response to either the whole of, or part of, a proposed musical composition.

Notes

It is to be understood that the above-referenced arrangements are only illustrative of the application for the principles of the present invention. Numerous modifications and alternative arrangements can be devised without departing from the spirit and scope of the present invention. While the present invention has been shown in the drawings and fully described above with particularity and detail in connection with what is presently deemed to be the most practical and preferred example(s) of the invention, it will be apparent to those of ordinary skill in the art that numerous modifications can be made without departing from the principles and concepts of the invention as set forth herein. For example, the mathematical equations given in this text are specific and non-limiting examples only.

Appendix 1

X-System Technical Outline: Component Overview

Fundamentally, the X-System is comprised of 3 components, two of which are software, and one which is hardware.

One piece of software (the "Music Analyser") is used in an offline (not directly linked to the real-time operation of the system) mode to analyse the candidate music files, and to build an estimation of their excitement/affect influence.

The second software part is the playback component. This is responsible for actually playing the music files, and also for receiving data from the sensor hardware, and using it to update its internal model which determines subsequent tracks to play.

Finally, the hardware component consists of a number of sensors which gather real-time data from the local environment, primarily from the actual user.

Detailed Descriptions

Music Analysis

The analysis aspect of the music analysis subsystem has been described in detail elsewhere, and is not covered here. This section covers only the integration aspects. As mentioned, this is expected to operate primarily in an offline, non-interactive fashion. It will be run periodically against a batch of music inputs, which will result in a set of values describing certain properties of the track. These values can also be combined to produce a single 'excitement' figure for the track, which is used by the playback system. The benefit of storing the components individually is that as data is gathered and used to tune system, excitement values can be recomputed with different coefficients without the need to re-analyse the entire track, greatly reducing overhead.

All outputs of the analysis will be stored in a database, indexed on a number of parameters, including at least track and artist identifiers, and some form of acoustic signature which is relatively tolerant of encoding differences or background noise.

These indexes will be used when a user 'imports' their music collection to the system. If any tracks already exist in the database, their values do not need to be recomputed.

The feedback process will be an opt-in system in which users agree to provide anonymised information about their usage of the system in order to improve it.

Automated features such as normalised change in arousal, replay/skip of suggested tracks, and aggregate sensor data can be used. Explicit feedback in the form of like/dislike acknowledgements, and occasional randomised questionnaires may also be used.

Use of feedback to guide system parameters may be on both a global and per-user basis. Large scale data mining, pattern recognition, machine learning systems will be used to improve affect/arousal estimation of music.

The analysis component will be operated as an internet accessible service, either in conjunction with some music streaming service to provide the audio, or purely as a control system operating with the users personal music collection.

Where fast & reliable internet service is available, significant fraction of the processing can be offloaded to the hosted X-system service. This allows more intensive processing than on a typical end-device, and also secures the analyser IP.

Additional Uses

Beyond the primary aim of 'Arousal Adjustment' -- facilitating relaxation or excitement -- there are other possible uses for the music analysis. It can be used to add an additional dimension to music discovery and navigation, by observing the effect of a large number of short music samples on a user, and then suggesting tracks or artists with similar characteristics. If the system has been used by someone for any reasonable time and has a well-adapted personal model, this initial step may be unnecessary. Similarity navigation of "Music like Artist/Album X" may also be possible based on features determined during track analysis.

Playback and Decision

The playback component handles 2 tasks. Controlling the music playback, and operating a real-time arousal analysis/entrainment model, based on sensor input. The component may be responsible for actually playing the music, or may be a control layer on top of an existing media player such as iTunes/Windows Media Player, etc. The arousal analysis model will be based on the X-system INRM model, using the pre-computed values from the Music Analysis component as a starting point. The user will select a desired outcome, and the sensors will be used to gauge progress towards that outcome of each track. Explicit overrides will permit the user to manually skip a particular track either once, or to permanently blacklist it to ensure it will never be chosen again for them. In addition to their effect, these overrides will feed the decision model.

The capabilities of the component will be somewhat dependent on the environment it is operating in. On relatively low-power devices such as phones and portable music players, it may operate in a less precise, less computationally intensive mode, or if possible, offload some processing to a remote service.

For laptop/desktop/tablet applications, a more sophisticated model may be used. For niche uses, it may operate in conjunction with a visualiser or video playback component to enhance the entrainment effect.

It is likely that many users will wish to use the system from multiple different hosts, for example both their phone and laptop. The player requires some method of synchronising and sharing model data between these systems. This may be best implemented through (or on top of) some internet service similar to Apple iCloud or Google gDrive. This would also provide the channel for presenting data to the analysis system for modelling/training.

Additional Uses, Comments

Given enough training, it maybe possible to develop a version of the X-System that can operate at some level with no sensor feedback. This is likely to be less effective than a well-instrumented setup, but there may be sufficient value to the user in avoiding the complications of sensor purchase, upkeep, and inconvenience of wearing. If this proves impossible or undesirable, it may be possible to obtain some feedback through sensors without direct user attachment, for example a accelerometer in the phone carried in their pocket, or GPS in the same indicating their location.

Sensor Hardware

Currently, the sensing part of the system uses two distinct sensors. One is a pulse oximeter, which is used to monitor heart-rate, and the other is a skin conductance sensor, which measures the electrical conductance (the inverse of resistance) of the skin.

Pulse Oximeter

The pulse oximeter operates on the principle of wavelength-dependent absorption of light by the (oxy-)haemoglobin in the bloodstream. By comparing absorption values at red and infra-red wavelengths, the relative proportion of oxygenated blood can be determined, leading to the 'blood oxygen saturation' (spO2) figure. Tracking this value at a relatively high frequency allows detection of the sudden change which indicates a pulse due to a heart-beat, and hence, heart-beat rate can be determined. Whilst very useful in medical contexts, blood oxygenation does not change significantly or at timescales useful to the X-System, and only heart-rate data is collected.

The current system uses a COTS sensor, the Nonin 3150 WristOx2 wireless pulse oximeter. This device uses a soft rubber fingertip clip to house the light emitter/detectors, which is typical for the type of sensor. Alternatives exist which use sensors clipping gently to the lobe of the ear, as well as other parts of the body. This device uses Bluetooth (with the standard and generic SPP - Serial Port Protocol) for data transmission.

Future implementations of this sensor are likely to use sensor locations more convenient and less intrusive than a fingertip. The reliability and accuracy of the sensor is strongly improved by using direct transmission absorption (that is, directing light through a relatively thin body-part such as a finger or ear-lobe), but devices do exist which can operate in reflective mode, allowing them to be placed almost anywhere, although areas with high blood vessel density, and relatively close to the surface of the skin are to be preferred. One good site which fits well with the x-system goals would be as part of a watch strap, with the sensor on the inside of the wrist, where the buckle lies on a typical watch-strap.

Skin Conductance

Skin Conductance, variously termed EDA (Electro-Dermal activity), GSR (Galvanic Skin Resistance), or just Skin Resistance/Conductance, is a measure of the ability of the skin to carry electrical current. For obvious safety reasons, the current must be kept very low, and strictly limited. Baseline skin conductivity depends on a multitude of factors specific to individuals and their local environment, but on short timescales, the primary influence is that of sweat. Sweat, essentially just water high in electrolytes, is a good conductor, and its presence lowers the effective resistance of the skin. As an aside, Conductance (measured in Siemens/mhos) is defined as the inverse of resistance (in ohms). By convention conductance is used when describing these systems, although conversion to resistances is trivial.

Sweating is influenced by a variety of factors, but we are most interested in the relation to the parasympathetic nervous system. Increased arousal is strongly correlated with increased sweating, and hence increased skin conductance. This effect is relatively fast, on the order of seconds. The areas of the body with the highest density of sweat glands -- the working surfaces of the hands and feet -- are the most effective pickup locations, but other locations are possible, with varying results. The wrist and outer forearm have been shown to provide adequate results [ref available]

Measuring skin conductance can be achieved in several ways. The current sensor uses a simple potential divider with a high-precision resistor as one leg, and 2 skin contacts applied to the user serve as the other leg. The central node is also connected to a buffered ADC for measurement.

Other designs exist, and some prototype work has been done on using a Wheatstone Bridge -- a particular circuit arrangement which allows highly precise differential measurements -- to improve accuracy and noise rejection.

An important aspect of this parameter is that the value can vary over several orders of magnitude. Dry skin, in a cold, dry environment, can have conductances in the micro-Siemen (Mega-ohm) range, and extremely sweaty skin can go down to hundreds of milli-Siemen (1-1000 Ohms). Accurate measurement across this wide range presents some significant challenges in sensor design.

The existing sensor, as mentioned, uses a relatively unsophisticated potential divider. This is sampled at around 50Hz by an Analogue-to-Digital Converter (ADC) integrated into the sensor microcontroller (MCU).

The particular MCU used at present is the Texas Instruments MSP430F2774. In addition to the ADC, this device contains an integrated programmable gain amplifier (PGA), which is used to magnify the signal from 1x to 16x. This provides an effective increase in precision of 4 bits to the existing 10-bit ADC. Preceding the amplifier is another integrated Op-Amp which is used in follower (unity-gain) mode, which acts to buffer the signal, and present a high-impedance load to the voltage divider, ensuring that the reading is not skewed due to significant current flowing through the sampling subsystem.

The ADC input is sampled at approximately 50Hz. If the measured value falls into one of the two regions near the top and bottom of its full measurement range, the gain of the PGA pre-amp is adjusted to raise it towards the centre of the measurement range. Immediately following this adjustment (after a short settling period required by the amplifier) another sample is taken. A hysteresis method is implemented at the edge of each region to minimise the possibility of 'flip-flopping' repeatedly between 2 amplifier gain levels and interfering with the timely gathering of values. In addition, the relatively high sampling rate (50Hz) compared to the transmission rate of approximately 2Hz leaves plenty of room for amplifier adjustments. The high sample-rate readings are averaged using a simple low-pass (FIR) filter with a cutoff of 10Hz.

Samples which fall into these border regions and result in an amplification change are discarded once this second sample completes. Software semaphores are used in the firmware to ensure the communication subsystem cannot access the sample buffer whilst it is in use or contains unreliable data.

If the reading falls into a buffer region but the pre-amp is already set to the maximum or minimum value possible, the reading is stored and transmitted, but marked with a flag indicating a potential saturation/clipping error.

The MCU is also connected to a wireless radio module, which it uses to communicate with a USB base-station. The wireless communications operate in the same unregulated frequency band as WiFi and Bluetooth, at 2.4GHz. They are however of much lower power and data-rate, and are designed to co-exist with these other devices nearby.

Higher level radio communications are handled using a slightly modified version of the SimpliciTI proprietary network protocol on the sensor device and base station. This allows multiple sensors to operate in range of one another while ensuring that data is received by the correct base-station. Base stations are implemented using a second MSP430, this time with a USB interface, and which uses the standard USB-Serial device-driver which is supported by practically all host devices and operating systems. Layered on top of the network protocol is the X-System sensor protocol, which exists mainly to facilitate transmission of sensor readings, provide debugging output, and allow selective enabling/disabling of sensors to save power. The update frequency of the sensors can also be adjusted.

The sensors are battery powered, with in-situ charging possible over USB. This permits fully wireless operation, and minimises any noise that could be present in external power supply lines.

Notes

The above section describes the existing implementations, but there are a number of additional features planned, but not yet deployed. These include both upgrades to the current sensing modalities, and also the incorporation of additional types of sensor. Upgrades include:

- Heart-rate:
    - Reflective IR Pulse Oximeter suitable for wrist-mounted sensing.
    - High frequency plethysmographic sampling for heart waveform & rhythm analysis, beyond a simple 'heart-rate' value.
- Skin Conductance:
    - Wheatstone Bridge based skin conductance pickup, with discrete or integrated precision instrumentation amplifiers.
    - More sophisticated digital filtering stage
    - Use of synchronised accelerometer attached to/near the skin contacts used to mark readings as suspicious due to contact-movement artifacts.

Additional modalities include:

- EEG type sensors or 'caps' for brainwave activity
- Electromyograph muscular tone/trigger rate
- Multi-point ECG for high-resolution heart waveform
- Breathing depth/rate
- Eye-tracking/Gaze/blink analysis

Future sources of data which are not yet viable, but which would benefit the system include: stress hormone (e.g. Cortisol) plasma concentration, neural triggering rate, regional brain activity.

The primary obstacle to be overcome in the development of sensors is convenience. If aimed at a mass market, few users will tolerate cumbersome cables or obstructions of their hands or senses, in comparison to, for example, a therapeutic or medical market. Consolidation of sensors into a single package such as a wrist-watch or headphone style appliance would be ideal. Other possibilities include flexible circuits integrated into clothing or footwear.

A sensor package should be capable of interoperability with as many host devices as is feasible. This may include smart-phones, feature-phones, tablets, portable music players, laptops, desktops, home hifi, and in-car audio. The most common interfaces are likely WiFi or Bluetooth, although support varies significantly across the range of hosts described.

Appendix 2

Modelling Human Neuro-physiological Response

The following papers, which are incorporated by reference, provide information on modelling the neuro-physiological response of humans.

Aragon D, Farris C, Byers JF
The effects of harp music in vascular and thoracic surgical patients
Alternative Therapies in Health and Medicine 2002 September-October; 8(5): 52-4, 56-60

Baumgartner T, Lutz K, Schmidt CF, Jancke L
The emotional power of music: how music enhances the feeling of affective pictures
Brain Research 2006 February; 1075 (1): 151-64

Bernardi L, Porta C, Sleight P
Cardiovascular, cerebrovascular and respiratory changes induced by different types of music in musicians and non-musicians: the importance of silence
Heart (British Cardiac Society) 2006 April; 92(4): 445-52

Blood AJ, Zatorre RJ
Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion
Proceedings of the National Academy of Sciences USA. 2001 September 25; 98(20): 11818-23

Brown S, Martinez MJ, Parsons LM
Passive music listening spontaneously engages limbic and paralimbic systems
Neuroreport 2004 September 15; 15(13): 2033-7

Brugge JF
Patterns of organisation in auditory cortex
Journal of the Acoustical Society of America 78(1/2) 1985 353-359

Byers JF, Smyth KA
Effect of a musical intervention on noise annoyance, heart rate, and blood pressure in cardiac surgery patients
American Journal of Critical Care 1997 May; 6(3): 183-91

Cardigan ME, Caruso NA, Haldeman SM, McNamara ME, Noyes DA, Spadafora MA, Carroll DL
The effects of music on cardiac patients on bed rest
Progress in Cardiovascular Nursing 2001 Winter; 16(1): 5-13

Decety J, Chaminade T
Neural correlates of feeling sympathy
Neuropsychologia 41 2003 127-138

Evers S, Suhr B
Changes of the neurotransmitter serotonin but not of hormones during short time music perception
European Archives of Psychiatry and Clinical Neuroscience 2000; 250(3): 144-7

Formisano E, Kim DS, Di Salle F, van de Moortele PF, Ugurbil K, Goebel R
Mirror-symmetric tonotopic maps in human primary auditory cortex
Neuron 40(4) 2003 859-869

Gallese V
The roots of empathy. The shared manifold hypothesis and the neural basis of intersubjectivity
Psychopathology, 36 2003 171-180

Gerra G, Zaimović A, Franchini D, Palladino M, Giucastro G, Reali N, Maestri D, Caccavari R, Delsignore R, Brambilla F
Neuroendocrine responses of healthy volunteers to 'techno-music': relationships with personality traits and emotional state
International Journal of Psychophysiology 1998 January; 28(1): 99-111

Grape C, Sandgren M, Hansson LO, Ericson M, Theorell T
Does singing promote well-being?: An empirical study of professional and amateur singers during a singing lesson
Integrative Physiological and Behavioral Science 2003 January-March; 38(1): 65-74

Griffiths TD, Buchel C, Frackowiak RS, Patterson RD
Analysis of temporal structure in sound by the human brain
Nature Neuroscience 1(5) 1998 422-427

Hebert S, Beland R, Dionne-Fournelle O, Crete M, Lupien SJ
Physiological stress response to video-game playing: the contribution of built-in music
Life Sciences 2005 April 1; 76(20): 2371-80

Holstege G, Bandler R, Saper CB (ed)
The emotional motor system
Progress in Brain Research 107, Elsevier, Amsterdam 1996

Iwanaga M
Relationship between heart rate and preference for tempo of music
Perceptual and Motor Skills 1995 October; 81(2): 435-40

Iwanaga M, Kobayashi A, Kawasaki C
Heart rate variability with repetitive exposure to music
Biological Psychology 2005 September; 70(1):61-6

Iwanaga M, Tsukamoto M
Effects of excitative and sedative music on subjective and physiological relaxation
Perceptual and Motor Skills 1997 August; 85(1): 287-96

Jeannerod M
Visual and action cuescontribute to the self-other distinction
Nature Neuroscience 7(3) 2004 421-422

Knight WE Rickard NS
Relaxing music prevents stress-induced increases in subjective anxiety, systolic blood pressure and heart rate in healthy males and females
Journal of Music Therapy 2001 Winter; 38(4): 254-72

Koelsch S, Fritz T, V Cramon DY, Muller K, Friederici AD
Investigating emotion with music: an fMRI study
Human Brain Mapping 2006 March; 27(3): 239-50

Kumar AM, Tims F, Cruess DG, Mintzer MJ, Ironson G, Loewenstein D, Cattan R, Fernandez JB, Eisdorfer C, Kumar M
Music therapy increases serum melatonin levels in patients with Alzheimer's disease
Alternative Therapies in Health and Medicine 1999 November; 5(6): 49-57

Lee DN
Guiding movement by coupling taus
Ecological Psychology 1998: 10(3-4): 221-250

Lee OK, Chung YF, Chan MF, Chan WM
Music and its effect on the physiological responses and anxiety levels of patients receiving mechanical ventilation: a pilot study
Journal of Clinical Nursing 2005 May; 14(5): 609-20

Li L, Korngut LM, Frost BJ, Beninger RJ
Prepulse inhibition following lesions of the inferior collicus: prepulse intensity functions
Physiology and Behavior 1998 August; 65(1): 133-9

McAdams S, Winsberg S, Donnadieu S, De Soete G, Krimphoff J
Perceptual scaling of synthesised musical timbres: common dimensions, specidicities, and latent subject classes
Psychological Research 58 1995 177-192

Meloni EG, Davis M
The dorsal cochlear nucleus contributes to a high intensity component of the acoustic

startle reflex in rats
Hearing Research 1998 May; 119(1-2): 69-80

Menon V, Levitin DJ, Smith BK, Lembke A, Krasnow RD, Glazer D, Glover GH, McAdams S
Neural correlates of timbre change in harmonic sounds
Neuroimage 17(4) 2002 1742-1754

Miluk-Kolasa B, Obminski Z, Stupnicki R, Golec L
Effects of music treatment on salivary cortisol in patients exposed to pre-surgical stress
Experimental and Clinical Endocrinology 1994; 102(2): 118-20

Mok E, Wong KY
Effects of music on patient anxiety
AORN Journal 2003 February; 77(2): 396-7, 401-6, 409-10

Molnar-Szakacs I, Overy K
Music and mirror neurons: from motion to 'e' motion
Social Cognitive Affective Neuroscience 1 2006 235-241

Nilsson U, Unosson M, Rawal N
Stress reduction and analgesia in patients exposed to calming music postoperatively: a randomized controlled trial
European Journal of Anaesthesiology 2005 February; 22(2): 96-102

Osborne N 1.
Music for children in zones of conflict and post-conflict
in Communicative Musicality ed. Malloch and Trevarthen OUP 2009

Osborne N 2.
Towards a chronobiology of musical rhythm
in Communicative Musicality ed. Malloch and Trevarthen OUP 2009

Osborne N 3.

Neuroscience and real world practice...
Annals of the New York Academy of Sciences 2012 (in publication)

Overy K, Molnar-Szakacs I
Being together in time: musical experience and the mirror neuron system
Music Perception 26 2009 489-504

Pachetti C, Aglieri R, Mancini F, Martignoni E, Nappi G
Active music therapy and Parkinson's disease: methods
Functional Neurology 1998 January-March;13(1): 57-67

Panksepp J, Trevarthen C
The neuroscience of emotion in music
in Communicative Musicality OUP 2009

Pantev C, Hoke M, Lehnertz K, Lutkenhoner B, Anogianakis G, Wittkowski W
Tonotopic organisation of the human auditory cortex revealed by transient auditory-evoked magnetic fields
Electroencephalographic Clinical Neurophysiology 69(2) 1988 160-170

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD
The processing of temporal pitch and melody information in the auditory cortes
Neuron 36(4) 2002 767-776

Penhune VB, Zatorre RJ, Feindel WH
The role of auditory cortex in retention of rhythmic patterns as studied in patients with temporal lobe removals including Heschl's gyrus
Neuropsychologia. 1999 March; 37(3):315-31

Peretz I
Listen to the brain: a biological perspective on musical emotions
in Juslin P, Sloboda J (eds) Music and Emotion: Theory and Research
OUP London 2001

Peretz I, Zatorre RJ (ed)
The cognitive neuroscience of music OUP 2003

Peretz I, Kolinsky R
Boundaries of separability between rhythm in music discrimination: a neuropsychological perspective
The Quarterly Journal of Experimental Psychology 1993 May; 46(2): 301-25

Reinhardt U
Investigations into synchronisation of heart rate and musical rhythm in relaxation therapy in patients with cancer pain (in German)
Forschende Komplementarmedizin 1999 June; 6(3): 135-41

Rencanzone GH, Schreiner CE, Merzenich MM
Plasticity in the frequency representations of primary auditory cortex following discrimination training in adult owl monkeys
Neuroscience 13(1) 1993 87-103

Rizzolati G, Fogassi I, Gallese V.
Neuro-physiological mechanisms underlying the understanding and imitation of action
Nature Reviews Neuroscience, 2, 2001 661-670

Schneider N, Schedlowski M, Schurmeyer TH, Becker H
Stress reduction through music in patients undergoing cerebral angiography
Neuroradiology 2001 June; 43(6): 472-6

Stefano GB, Zhu W, Cadet P, Salamon E, Mantione KJ
Music alters constitutively expressed opiate and cytokine processes in listeners
Medical Science Monitor 2004 Jun; 10(6): MS18-27

Sutoo D, Akiyama K
Music improves dopaminergic neurotransmission: demonstration based on the effect of music on blood pressure regulation
Brain Research 2004 August 6;1016(2): 255-62

Talavage TM, Sereno MIO, Melcher JR, Ledden PJ, Rosen BR, Dale AM
Tonotopic organisation in human auditory cortex revealed by progressions of frequency sensitivity
Journal of Neurophysiology 91(3) 2004 1282-1296

Trevarthen C
Musicality and the Intrinsic Motive Pulse: Evidence from human psychobiology and infant communication
Special Issue of Musicae Scientiae: Rhythm, Narrative and Origins of Human Communication 1999: 157-213

Trevarthen C, Malloch SN,
The Dance of Wellbeing: Defining the Musical Therapeutic Effect
Nordic Journal of Music Therapy 2000; 9(2): 65-126

Turner R, Ioannides AA
Brain, music and musicality: inferences from neuroimaging
in Communicative Musicality OUP 2009

Uedo N, Ishikawa H, Morimoto K, Ishihara R, Narahara H, Akedo I, Ioka T, Kaji I, Fukuda S
Reduction in salivary cortisol level by music therapy during colonoscopic examination
Hepato-gastroenterology 2004 March-April; 51(56): 451-3

Updike PA, Charles DM
Music Rx: physiological and emotional responses to taped music programs of preoperative patients awaiting plastic surgery
Annals of Plastic Surgery. 1987 July; 19(1): 29-33

Urakawa K, Yokoyama K
Music can enhance exercise-induced sympathetic dominancy assessed by heart rate variability

The Tohoku Journal of Experimental Medicine 2005 July; 206(3): 213-8

VanderArk SD, Ely D
Cortisol, biochemical, and galvanic skin responses to music stimuli of different preference values by college students in biology and music
Perceptual and Motor Skills. 1993 August; 77(1): 227-34

Warren JD, Uppenkamp S, Patterson RD, Griffiths TD
Separating pitch chroma and pitch height in the human brain
Proceedings of the National Academy of Sciences USA, 100(17) 2003 10038-10042

Wieser HG, Mazzola G
Musical consonances and dissonances: are they distinguished independently by the right and left hippocampi?
Neuropsychologia 1986; 24(6): 805-12

Yamamoto T, Ohkuwa T, Itoh H, Kitoh M, Terasawa J, Tsuda T, Kitagawa S, Sato Y
Effects of pre-exercise listening to slow and fast rhythm music on supramaximal cycle performance and selected metabolic variables
Archives of Physiology and Biochemistry 2003 July; 111(3): 211-4

Zatorre RJ, Peretz I (ed)
The biological foundations of music
New York Academy of Sciences 2001

Zatorre R,J, Evans AC, Meyer E
Neural mechanisms underlying melodic perception and memory for pitch
Journal of Neuroscience 14(4) 1994 1908-1919

CLAIMS

1. Computer implemented system for analysing sounds, such as audio tracks, the system automatically analysing sounds according to musical parameters derived from or associated with a predictive model of the neuro-physiological functioning and response to sounds by one or more of the human lower cortical, limbic and subcortical regions in the brain;

and in which the system analyses sounds so that appropriate sounds can be selected and played to a listener in order to stimulate and/or manipulate neuro-physiological arousal in that listener.

2. System of Claim 1 in which the system is adapted to automatically analyse sounds and store the results of that analysis in a database so that appropriate sounds can subsequently be selected from that database and played to a listener to provide a desired stimulation and/or manipulation of neuro-physiological arousal in that listener.

3. System of Claim 1 or 2 where the musical parameters relate to rhythmicity.

4. System of any preceding Claim where the musical parameters relate to harmonicity, being the degree of correspondence to the harmonic series.

5. System of any preceding Claim where the musical parameters relate to turbulence, being a measure of rate of change and extent of change in musical experience.

6. System of Claim 3 and any Claim dependent on Claim 3, which predictively models primitive spinal pathways and the pre-motor loop (such as the basal ganglia, vestibular system, cerebellum), all concerned with primal responses to rhythmic impulses, by analysing beat induction, using a specifically calibrated onset window.

7. System of Claim 3 and any Claim dependent on Claim 3, which predictively models rhythmic pattern recognition and retention regions (such as the secondary auditory cortex of the temporal lobes) by using self-similarity/auto-correlation algorithms.

8. System of Claim 3 and any Claim dependent on Claim 3 which predictively models the activation of mirror neuron systems, which detect the power, trajectory and intentionality of rhythmic activity, through one or more of: indices of rhythmic power, including computation of volume levels, volume peak density, "troughs", or the absence of energy and, dynamic profiles of performance energy.

9. System of Claim 3 and any Claim dependent on Claim 3 which predictively models activation of mirror neuron systems by analysing the profile of expenditure of energy (precipitous for high arousal, smooth for low) before and in between onsets, important mirror neuron information, by a computation of profiles of energy flow leading to significant articulations.

10. System of Claim 4 and any Claim dependent on Claim 4 which predictively models the functioning and response of Heschl's Gyrus to sound by determining levels of harmonicity and inharmonicity.

11. System of Claim 4 and any Claim dependent on Claim 4 which detects a principal fundamental through calculation of the harmonic product spectrum, then establishes degrees of harmonicity both within and among the spectra of different fundamentals.

12. System of Claim 11, in which the detection of a principal fundamental and the establishment of degrees of harmonicity is applied both "vertically" to instantaneous moments, and "horizontally" to progressions of pitches and spectra in time (related to the tonotopic mapping of the area around Heschl's Gyrus) and expressed in terms of linear harmonic cost, which represents both the rate at which the fundamental is changing, and the harmonic distance of the changes.

13. System of Claim 12 which predictively models the neurophysiological sensing of simple timbre by Heschl's gyrus, superior temporal sulcus, circular insular sulcus by analysing windows of vertical harmonicity at instantaneous moments.

14. System of Claim 12 which predictively models melodic and harmonic progressions in terms of how far each STFT time slices deviates from the simple ratios of the harmonic series: Linear harmonic cost arises from STFT time slices whose fundamental frequency

differs from that of the previous slice; Time slices with no change in fundamental have a cost of zero.

15. System of Claim 5 and any claim dependent on Claim 5, in which turbulence combines indices of change in rhythmicity and harmonicity, with auditory brainstem and cortical activity innervating the amygdala, hippocampus and core emotional regions affecting neurotransmission and endocrine systems, including the HPA axis, dopamine circuits and levels of, for example, norepinephrine, melatonin and oxytocin.

16. System of any preceding Claim where the analysis of sounds operates in real-time on locally stored music data and the system includes software, firmware and/or hardware running on a personal computing device.

17. System of any preceding Claim where the analysis of sounds operates in real-time on music data stored on a server and the system includes software, firmware and/or hardware running on that server or an associated server.

18. System of any preceding Claim in which data defining the rhythmicity, harmonicity and turbulence for a particular sound, such as a music track or portion of that track, are combined in such a manner as to give a single output representing arousal

19. System of Claim 3 and any preceding Claim dependent on Claim 3 that determines rhythmicity using an equation that relates R to B and S, such as the equation $R=\sqrt{B} * S^2$, and where R is rhythmicity, B is beats per minute, S is the mean of the beat strength.

20. System of Claim 3 and any preceding Claim dependent on Claim 3 that determines rhythmicity using an equation that relates R to B and S, such as the equation $I = C/10 - H$, where I is inharmonicity, C is linear harmonic cost and H is instantaneous harmonicity.

21. System of Claim 5 and any preceding Claim dependent on Claim 5 that determines turbulence using an equation that links T to H and P, such as $T=dH/dt * P$, where T is turbulence, H is harmonicity and P is energy during peak volume.

22. System of any preceding Claim in which the values of rhythmicty, harmonicity and, if applicable, turbulence for a given track are combined and mapped to an n dimensional point p characterising physiological state E.

23. System of Claim 22 that determines excitement using an equation that links E to I, R and T, such as where E=(10*I*R)+T, where E is the excitement state, R is the rhythmicity, I is the inharmonicity and T is the turbulence.

24. System of Claim 22 in which the variations in the values of rhythmicity, harmonicity and, if applicable, turbulence for a given track are mapped to an n dimensional point p characterising the directional effect on physiological state.

25. System of any of any preceding Claim, wherein the sound is analysed for the purpose of navigation, or discovery, or retrieval, or selection, or matching for a specific requirement, or for playlist creation, or for entraining affect or for entraining mood.

26. System of any preceding Claims, wherein the sound is any of: music; the ambient or background noise in a workplace, cinema, home, shop, vehicle, car, train, aircraft; sounds of nature (wind, ocean etc.), sounds of animals, sonifications (planets, stars, flowers, trees, financial markets, cell activity etc.).

27. System of any preceding Claim, wherein the system is for selecting music tracks for playback to a human subject according to a preselected desired arousal state of the human subject, the music tracks being selected according to the model of human neuro-physiological response to music, the neuro-physiological model being used to select the music tracks for playback according to the neuro-physiological response to the music tracks predicted by the neuro-physiological model.

28. System of any preceding Claim, wherein the system comprises a computer operable to select the music tracks for playback to the human subject according to the preselected desired (the arousal state of the human subject, the music tracks being selected according to the model of human neuro-physiological response to music, the model running on the computer, the neuro-physiological model being used by the computer to select the music

tracks for playback according to the neuro-physiological response to the music tracks predicted by the neuro-physiological model.

29. System of any preceding Claim, wherein the computer is operable to receive a selection of the desired arousal state of the human subject.

30. System of any preceding Claim, wherein the human subject both selects their desired arousal state and is the recipient of the music track output generated by the system.

31. System of any of Claims 27 - 30, wherein selection of the desired arousal state is made by a party who is not the human subject.

32. System of any preceding Claim, wherein a user interface is presented to a user such that the user interface is operable to enable the user to choose from a menu of activities in order for the system to establish a target level of arousal and affect that will facilitate the chosen activity.

33. System of any preceding Claim, wherein a primary music library and analysis software resides on a computer operable, remotely or locally, by a human subject, with the ability to transfer a selection of music to a personal music player device, which then generates a dynamic playlist based on the available music.

34. System of any preceding Claim, wherein the model of human neuro-physiological response to audio tracks is refined through machine learning, making use of linear regressive and/or neural network approaches.

35. System of any preceding Claim, wherein the model of human neuro-physiological response to audio tracks is refined through machine learning processes and statistical analysis.

36. System of any previous Claim, wherein bio-metric data of a human subject is measured using a sensor to determine neuro-physiological arousal of the human subject.

37. System of any previous Claim, wherein the system comprises a biofeedback system including a sensor.

38. System of Claim 37, wherein the biofeedback system comprises the sensor, a processor and a data repository storing music categorisation data, the sensor being arranged to measure one or more parameters on a human subject and communicate said measurements to the processor, wherein the processor is arranged to select a music piece in dependence on the measurements and on the music categorisation data.

39. System of any preceding Claim, wherein a sensor is used to measure the state of arousal of a human subject, and music categorised by predictive modelling of the model of human neuro-physiological response is streamed or otherwise provided in order to achieve the preselected desired arousal state of the human subject.

40. System of any previous Claim, wherein the system comprises a sensor such that once the sensor is activated the system measures a human subject's initial level of neuro-physiological arousal and the system automatically constructs a playlist that will first mirror this level of arousal, then direct the human subject towards, and help to maintain them at, the preselected desired arousal state of the human subject.

41. System of any of Claims 36 - 40, wherein the sensor is in the form of a wristband, or of adapted headphones, or of Electroencephalography (EEG) 'caps', or of gloves.

42. System of any of Claims 36 - 41, wherein the sensor captures data comprising biometric parameters including at least one of: heart rate, heart rate including pulse rhythm analysis, blood pressure, adrenaline and oxytocin levels, muscular tension, brain waves and galvanic skin conductivity.

43. System of any of Claims 36 - 42, wherein the system includes a sensor and a software package.

44. System of any of Claims 1 - 35, the system comprising software but no sensor, the system reliant on average expected response.

45. System of any preceding Claim, where a playlist is created in order to entrain or maintain arousal and direct state of mind and/or affect.

46. System of any previous Claim, wherein measurement of a level of neuro-physiological arousal of an individual is automated and is expressed as a value in order to enable it to be mirrored by a musical effect of any one of a theoretically unlimited number of pieces of music in a database.

47. System as claimed in any preceding claim operable to detect emotional arousal parameters information of an audio track and further operable to embed the information into the audio track or into an electronic link to the audio track or as metadata associated with the track.

48. System as claimed in any preceding claim operable to enable an automated search of music stored on remote or local databases for music with signatures meeting defined criteria.

49. System as claimed in any preceding claim operable to share arousal values in a social networking application.

50. System as claimed in any preceding claim operable to share arousal values (a) associated with a user interacting with a search engine such as Google®, with that search engine, so that the search engine can then use those values to optimise search and/or advertisement selection by the search engine

51. System as claimed in any preceding claim operable to share arousal values associated with a user browsing a specific website or pages in a website with a website optimisation system so that the website optimisation system can use those values to optimise the website and/or specific pages.

52. Method for analysing audio tracks, comprising the step of the audio tracks being automatically analysed according to a predictive model of human neuro-physiological functioning and response to sounds by one or more of the human lower cortical, limbic and subcortical regions in the brain.

53. Method of Claim 52, including such further steps as needed to enable a system for analysing audio tracks to operate as a system as claimed in any previous Claim 1 – 51.

54. Method for analysing audio tracks for playback to a human subject according to a preselected desired arousal state of the human subject, comprising the steps of:
(i) storing a set of individual audio tracks operable for selection for playback;
(ii) predicting a neuro-physiological response to the individual audio tracks according to a neuro-physiological model of the functioning and response of one or more of the human lower cortical, limbic and subcortical regions in the brain to sounds;
(iii) receiving a selected desired arousal state of the human subject, and
(iv) selecting audio tracks according to the predicted neuro-physiological response to the individual music tracks, and according to the selected desired arousal state of the human subject.

55. Method of any of Claims 52 - 54, wherein the audio tracks are analysed for the purpose of navigation, or discovery, or retrieval, or selection, or matching for a specific requirement, or for playlist creation, or for entraining affect or for entraining mood.

56. Method of any of Claims 52 - 55, wherein a user interface is presented to a user, further comprising the steps of:
(i) the user choosing from a menu of activities in the user interface
(ii) the system establishing a target level of arousal and affect that will facilitate the chosen activity.

57. Method of any of Claims 52 - 56, further comprising the step of an automated categorization process classifying music tracks and indexing them according to values expressed in a Musical Effect Matrix M.

58. Method of Claim 52 - 57 further comprising the step of analysing tracks for their universal musical values of rhythmicity and inharmonicity.

59. Method of Claim 52 - 58 further comprising the step of analysing tracks for their universal musical values of turbulence.

60. Method of Claim 52 - 59, wherein values of rhythmicity, inharmonicity and, if applicable, turbulence are automatically determined using signal processing techniques.

61. Method of Claims 52 - 60, further comprising the step of combining values of rhythmicity, inharmonicity and turbulence to yield a measure of excitement or arousal.

62. Method of Claim 61, wherein excitement E equals (10* inharmonicity I* rhythmicity R)+ turbulence T.

63. Computer program product operable when running on a computer to analyse audio tracks for playback to a human subject, the computer program product applying an algorithm based on a predictive model of human neuro-physiological functioning and response to the audio tracks by one or more of the human lower cortical, limbic and subcortical regions in the brain.

64. Computer program product of Claim 63, wherein the predictive model of human neuro-physiological response is a deterministic model of human neuro-physiological response.

65. Computer program product operable when running on a computer to analyse audio tracks for playback to a human subject according to a preselected desired arousal state of the human subject, the computer program product operable to perform the steps of:
(i) identifying a set of individual audio tracks operable for selection for playback;
(ii) predicting a neuro-physiological response to the individual audio tracks according to a neuro-physiological model;
(iii) receiving a selected desired arousal state of the human subject;
(iv) selecting for playback audio tracks according to the predicted neuro-physiological response to the individual audio tracks, and according to the selected desired arousal state of the human subject.

66. A computer-implemented method of categorizing sound (such as any piece of music regardless of genre or cultural origin) according to musical parameters derived from

a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to the pieces of music in such a way that it may be selected (e.g. automatically based on biometric data captured by a sensor) to entrain neuro-physiological arousal towards a target level;

67. System as claimed in any previous System Claim 1 - 51 adapted to provide the automatic categorisation of sound (such as pieces of music) in a remote database (e.g. according to musical parameters derived from a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to the pieces of music).

68. An automated diagnosis of the level of lower cortical, limbic and subcortical neuro-physiological arousal of an individual and expressing it as a value in order to correspond to the musical effect of any one of a theoretically unlimited number of pieces of music in a database.

69. A computer-implemented method of creating a playlist of tracks generated by automatically or manually analysing musical parameters derived from a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to the pieces of music in order to entrain arousal and direct state of mind and/or affect.

70. The computer implemented method of claim 69 including the further steps of:

a) choosing a subset of the music in a database by reference to existing descriptive metadata, if available, such as genre or user-created playlist;

b) selecting from this subset of music a number of pieces that will correspond to the user's initial level of lower cortical, limbic and subcortical neuro-physiological arousal by matching it to music contained in the relevant row of the musical effect matrix ;

c) selecting a target state of mind and/or affect;

d) selecting a series of ascending or descending musical effect values which correspond to the expected entrainment path from the initial to the required level of neuro-physiological arousal;

e) on the basis of this series of values, selecting qualified content from the music database;

f) choosing at random a playlist from the qualified content subject to other rules such as genre preference, an anti-repetition rule or the Unites States' Digital Millennium Copyright Act (DMCA) rules;

g) repeating the calculation of the playlist at intervals, based on continual or regular or occasional biometric feedback.

71. A method of determining the sufficiency of a (e.g. personal) database of music for the entrainment of affect using the system of any preceding system Claim 1 to 51 and of then displaying information to the user with regard to sufficiency or insufficiency.

72. A method of recommending a complement of musical content for a personal database of music in order to ensure sufficiency, by using musical parameters derived from a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to that music.

73. A method of selecting music that has a similar musical effect, (e.g. according to musical parameters derived from a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to the pieces of music).

74. A method of categorising music according to its musical effect rather than its descriptive attributes, using a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to that music.

75. A method of ordering a series of pieces of music in a playlist by matching the musical effect of each piece with a temporal series of values described by a musical effect vector, derived from a predictive model of human lower cortical, limbic and subcortical neuro-physiological functioning and response to that music.

76. A method of manipulating the arousal of a user by using any of the above methods or systems.

77. A method to modify the properties of ambient sound in any given environment, in order to produce a desired neuro-physiological response in the listener, by using any of the above methods or systems.

78. A system adapted to perform any of the above methods.

79. Software, firmware, SoCs or audio stacks programmed or adapted to perform any of the above methods or to form part of the system defined in any of the above systems.

80. A computing device, such as a smartphone or tablet, adapted to manipulate the arousal of a user by using any of the above methods or by using or including any of the above systems, software, firmware, SoCs or audio stacks.

81. Sensors adapted to work with the computing device defined in Claim 80.

Innate Neurophysiological Response to Music (Osborne 2009)

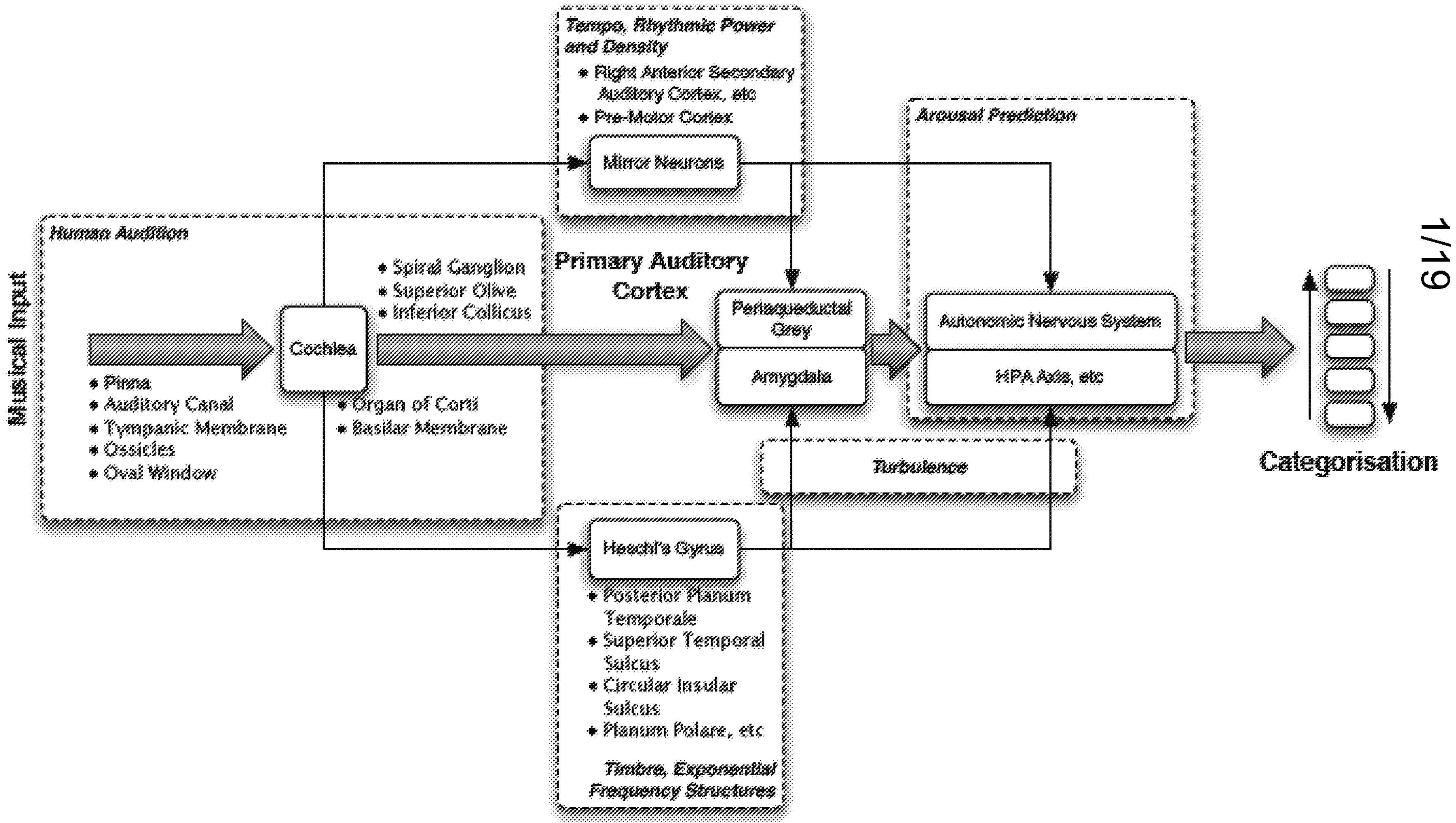


FIGURE 1

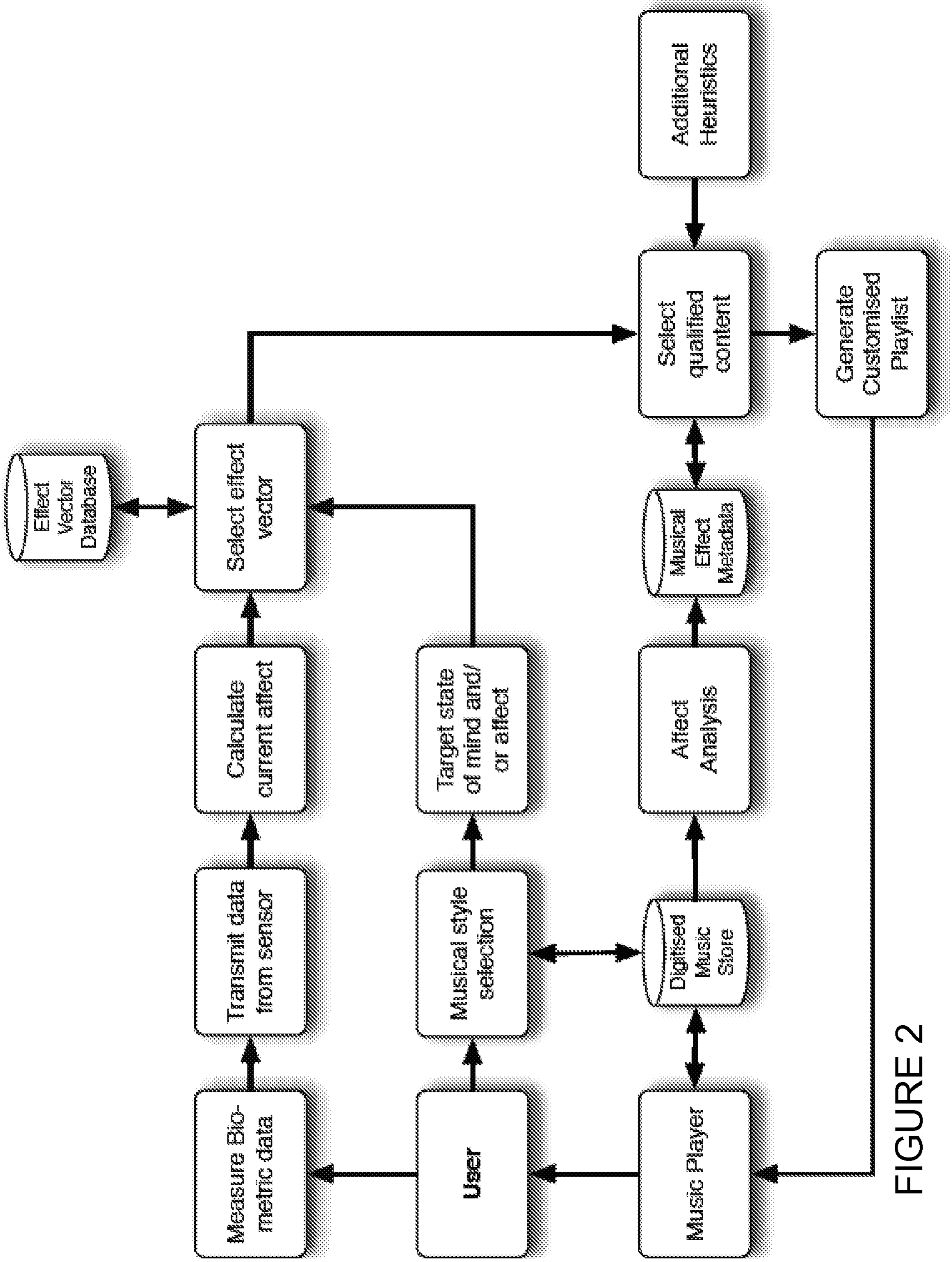

Measure Bio-metric data
Transmit data from sensor
Calculate current affect
Select effect vector
Effect Vector Database
User
Musical style selection
Target state of mind and/ or affect
Music Player
Digitised Music Store
Affect Analysis
Musical Effect Metadata
Select qualified content
Additional Heuristics
Generate Customised Playlist

FIGURE 2

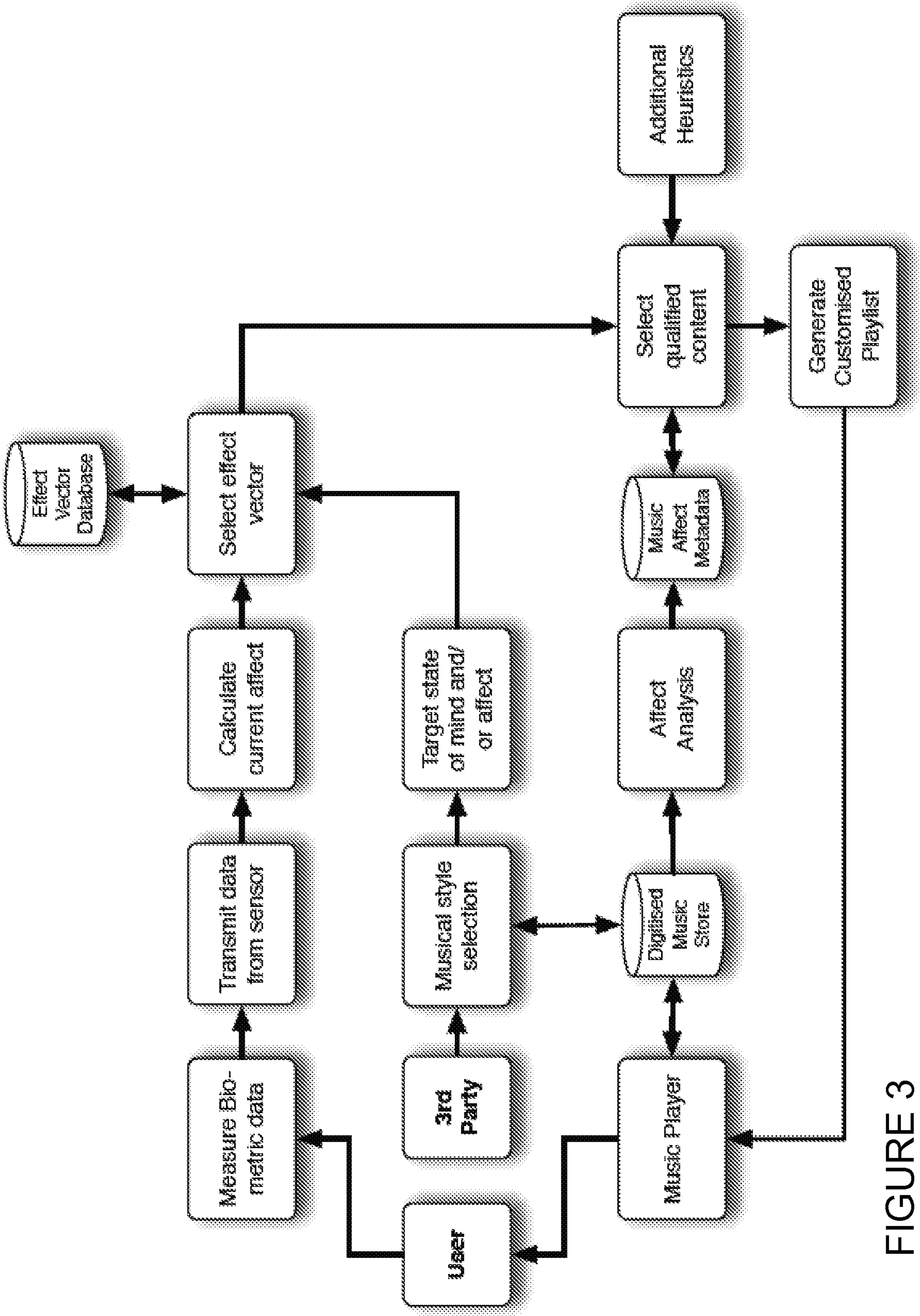

Effect Vector Database
Select effect vector
Calculate current affect
Transmit data from sensor
Measure Bio-metric data
Target state of mind and/ or affect
Musical style selection
3rd Party
User
Additional Heuristics
Select qualified content
Music Affect Metadata
Affect Analysis
Digitised Music Store
Music Player
Generate Customised Playlist

FIGURE 3

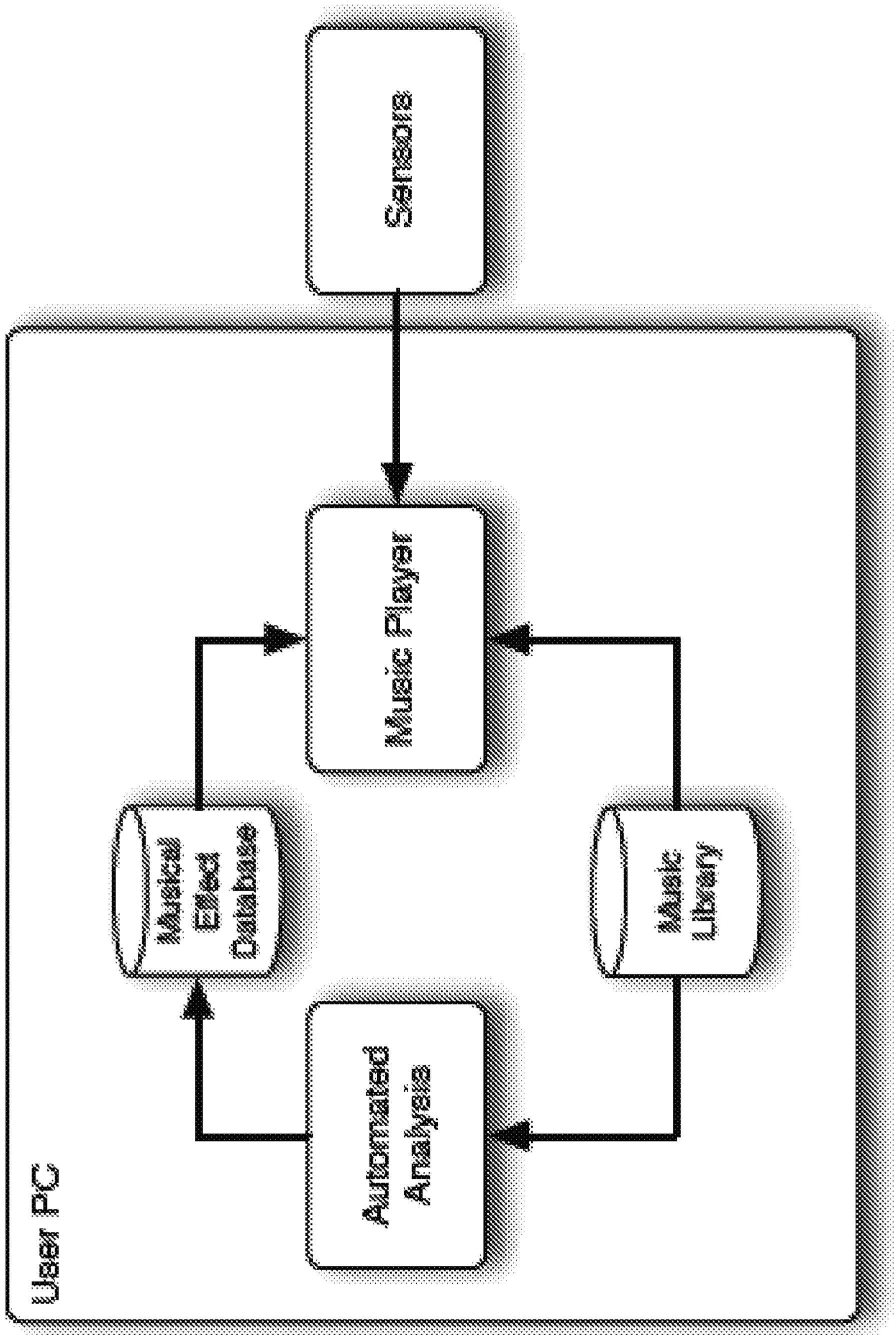

Sensors
User PC
Music Player
Musical Effect Database
Music Library
Automated Analysis

FIGURE 4

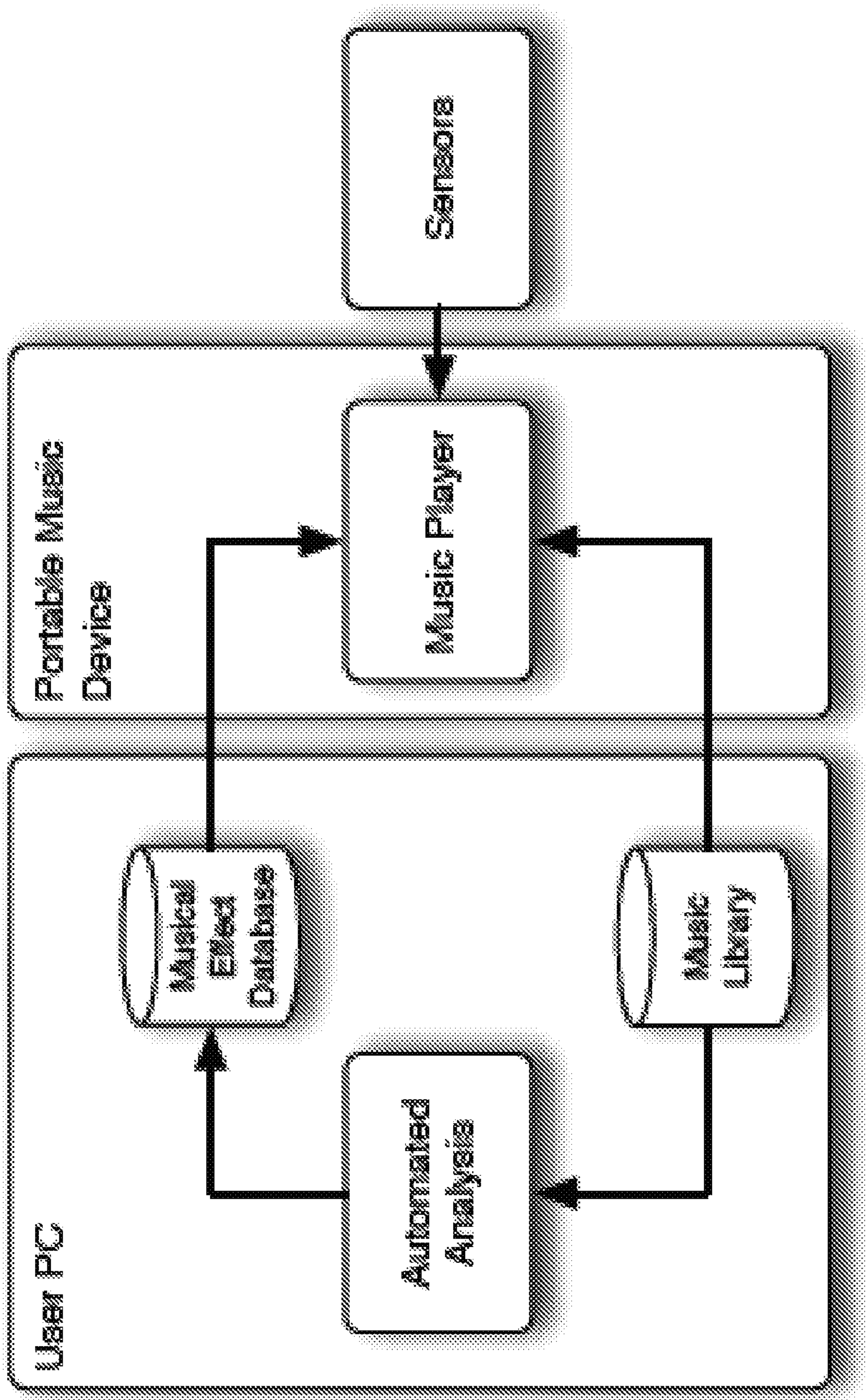

Sensors
Portable Music Device
Music Player
User PC
Musical Effect Database
Music Library
Automated Analysis

FIGURE 5

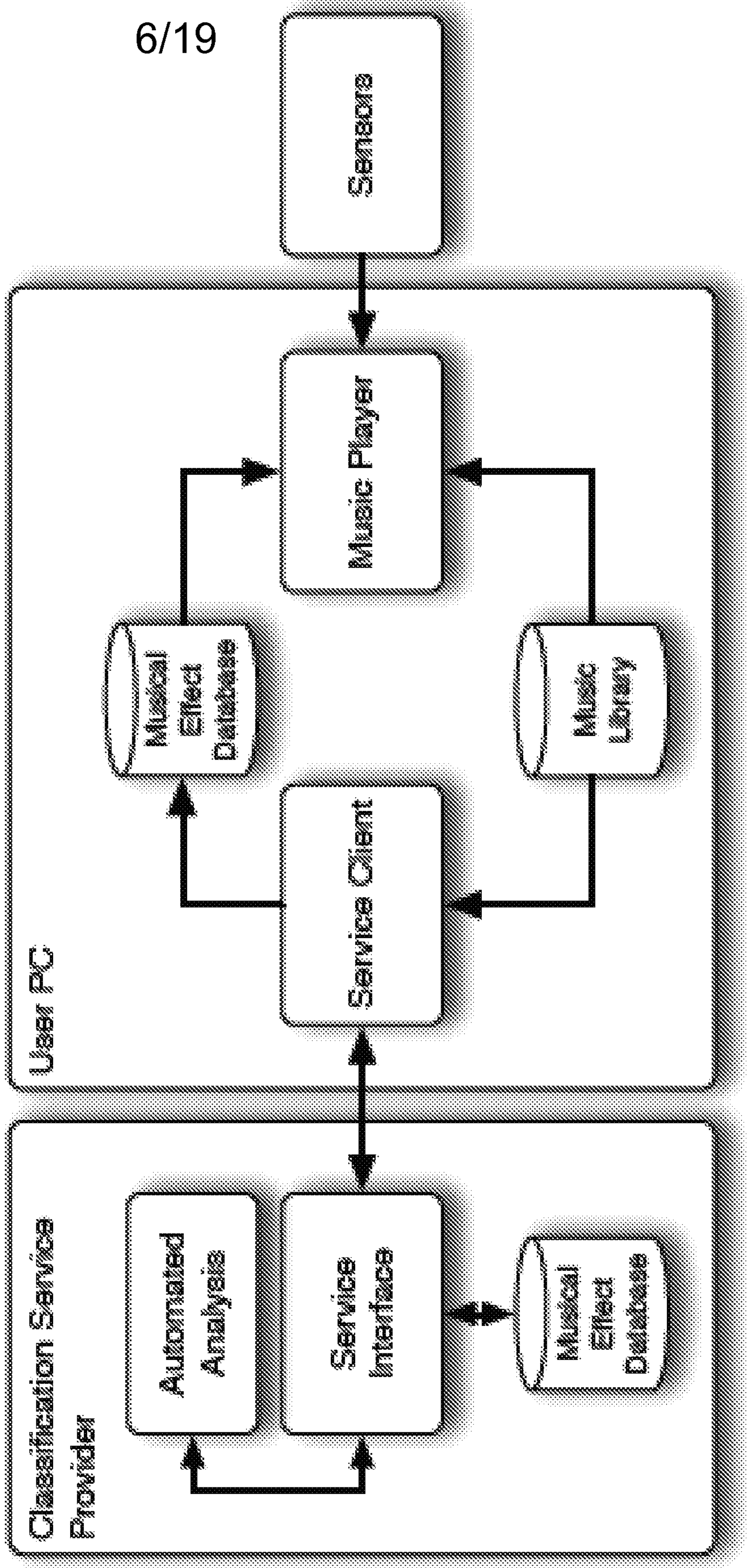
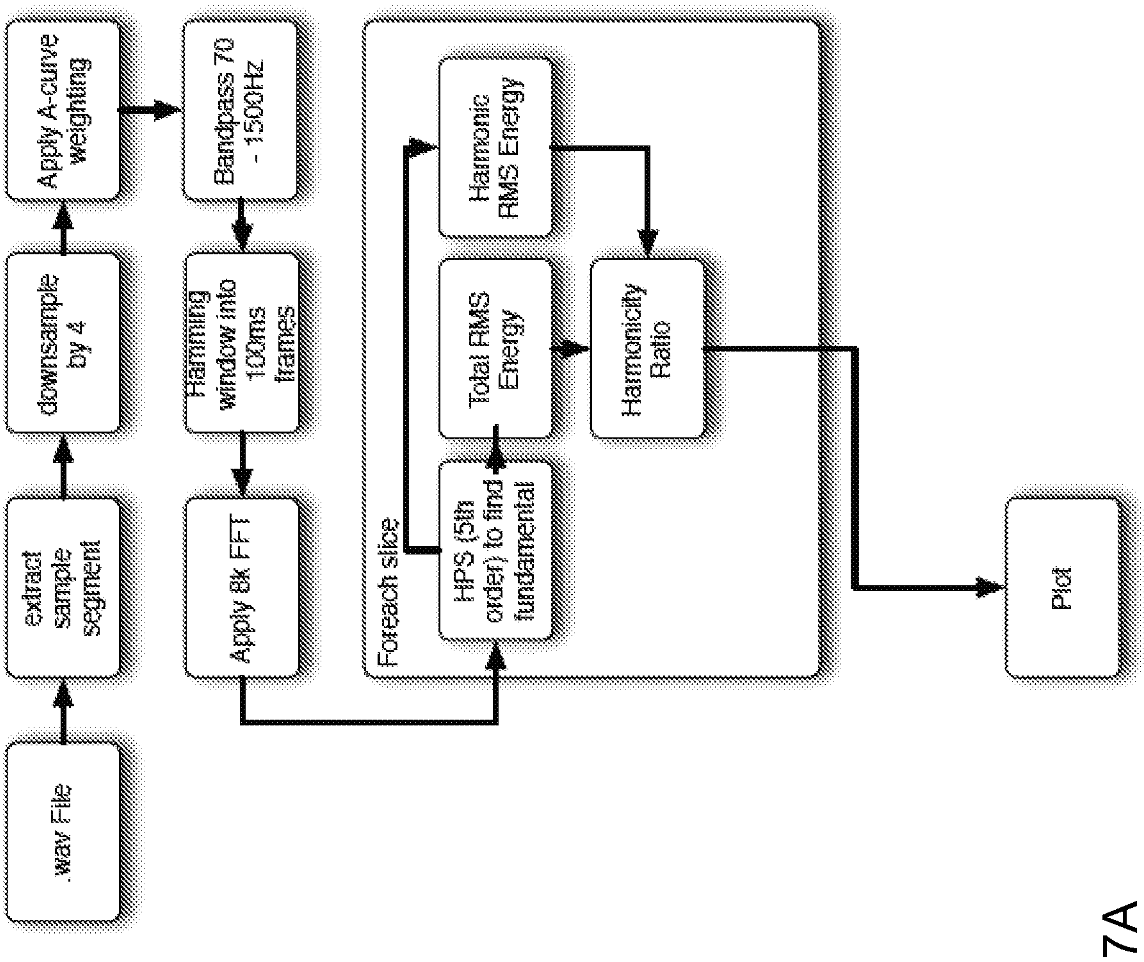

Sensors
User PC
Music Player
Musical Effect Database
Music Library
Service Client
Classification Service Provider
Automated Analysis
Service Interface
Musical Effect Database

FIGURE 6



.wav File
extract sample segment
downsample by 4
Apply A-curve weighting
Bandpass 70 - 1500Hz
Hamming window into 100ms frames
Apply 8k FFT
Foreach slice
HPS (5th order) to find fundamental
Total RMS Energy
Harmonic RMS Energy
Harmonicity Ratio
Plot

FIGURE 7A

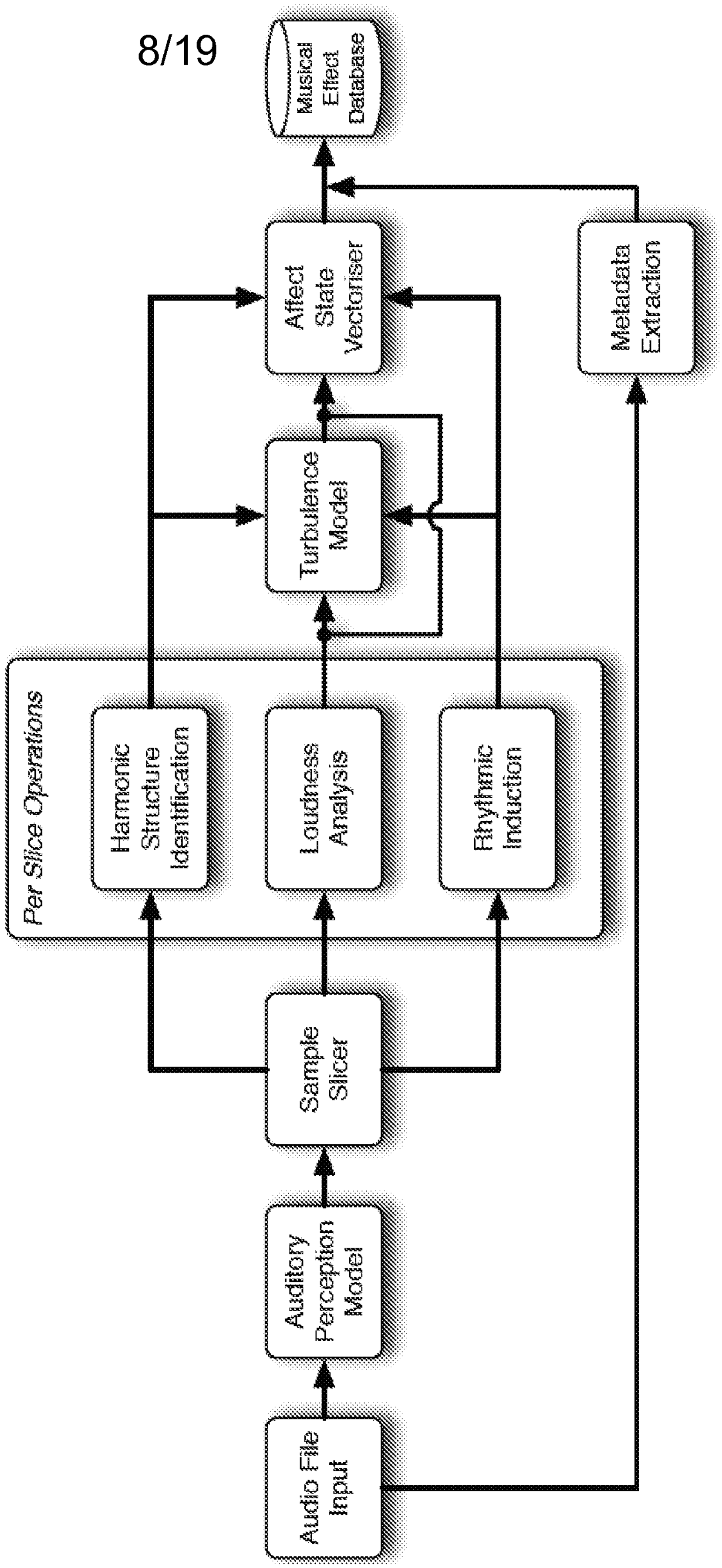

Audio File Input
Auditory Perception Model
Sample Slicer
Per Slice Operations
Harmonic Structure Identification
Loudness Analysis
Rhythmic Induction
Turbulence Model
Affect State Vectoriser
Musical Effect Database
Metadata Extraction

FIGURE 7B

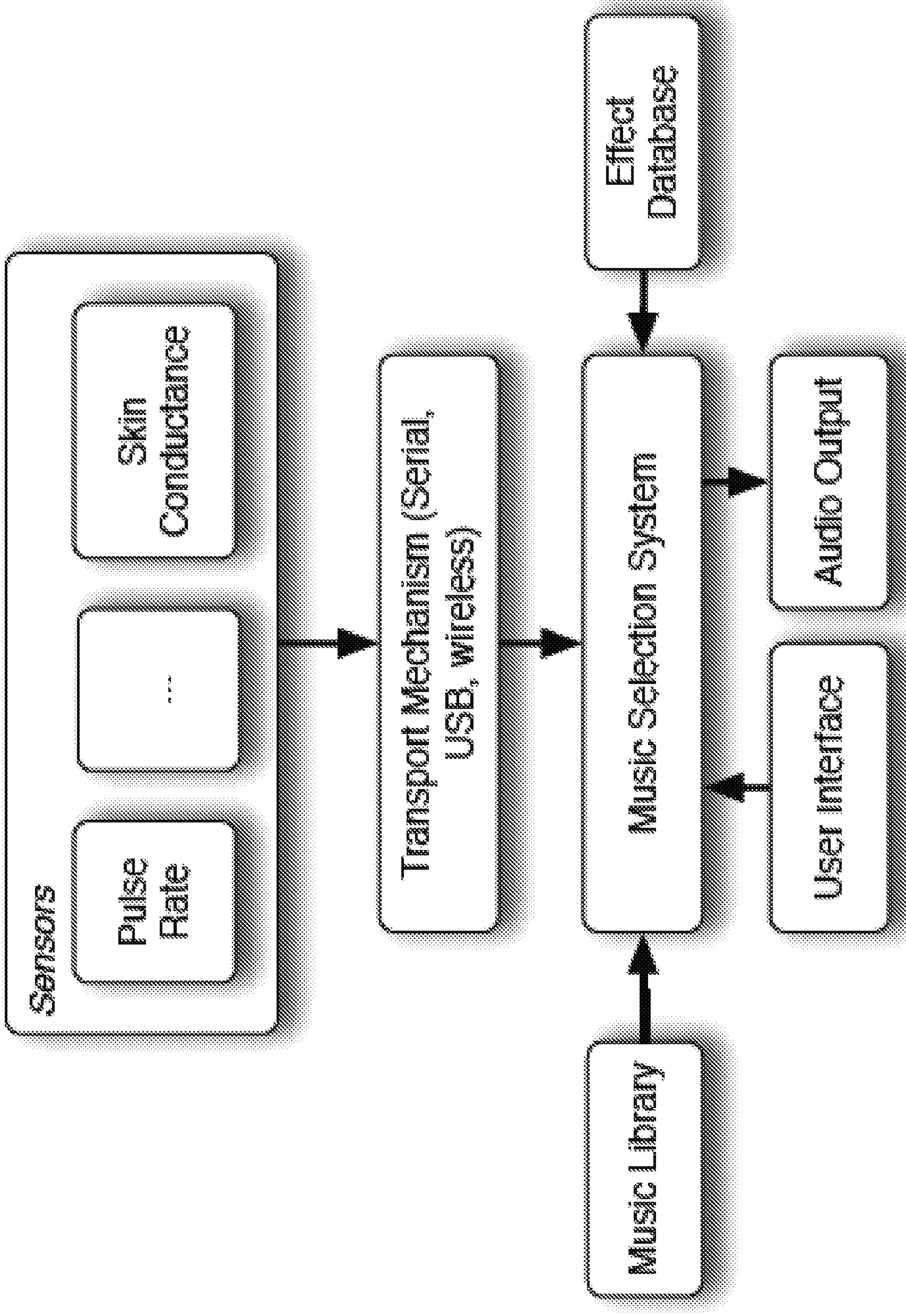

Sensors
Pulse Rate
...
Skin Conductance
Transport Mechanism (Serial, USB, wireless)
Effect Database
Music Selection System
Music Library
Audio Output
User Interface

FIGURE 8

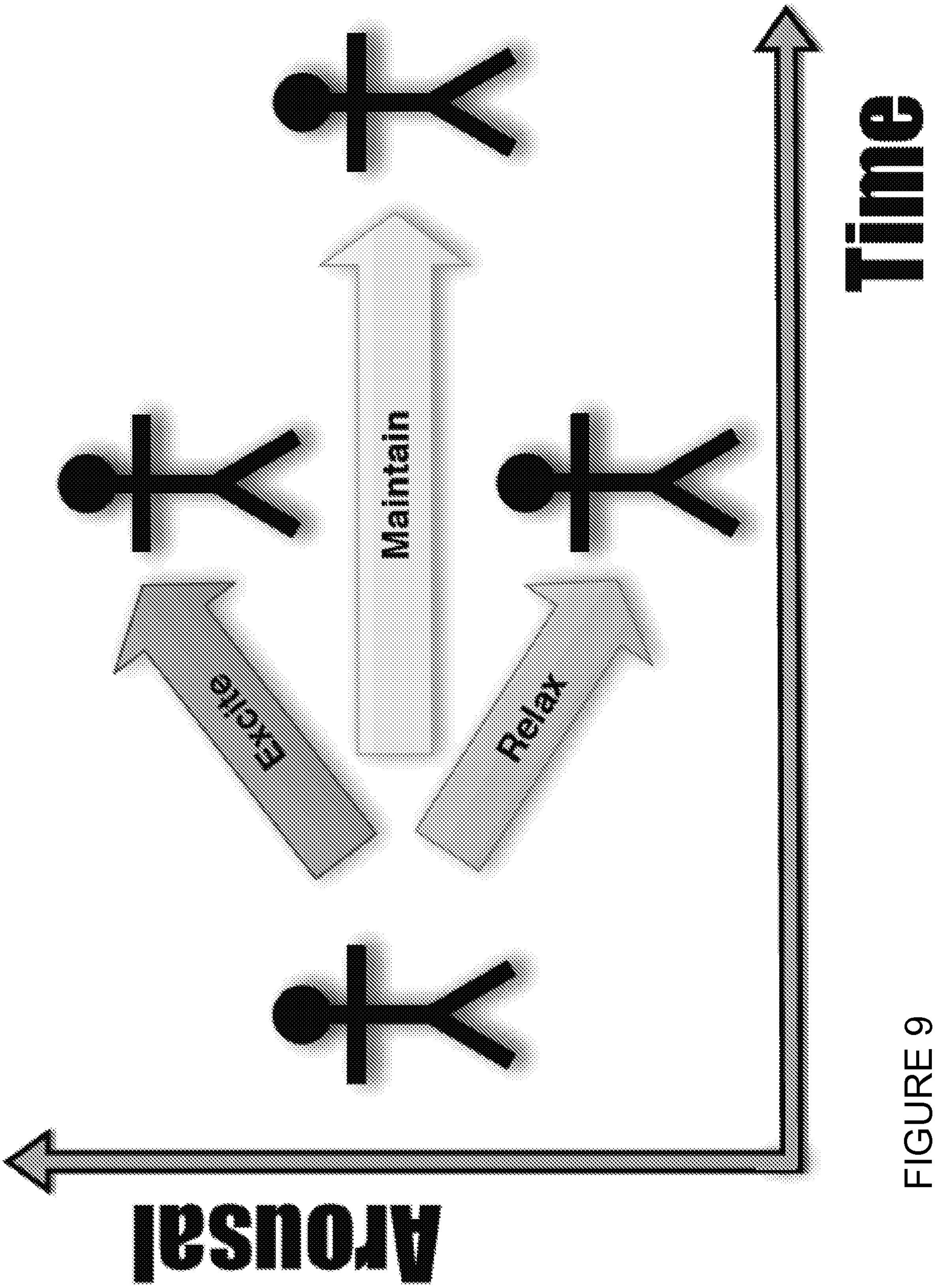


FIGURE 9

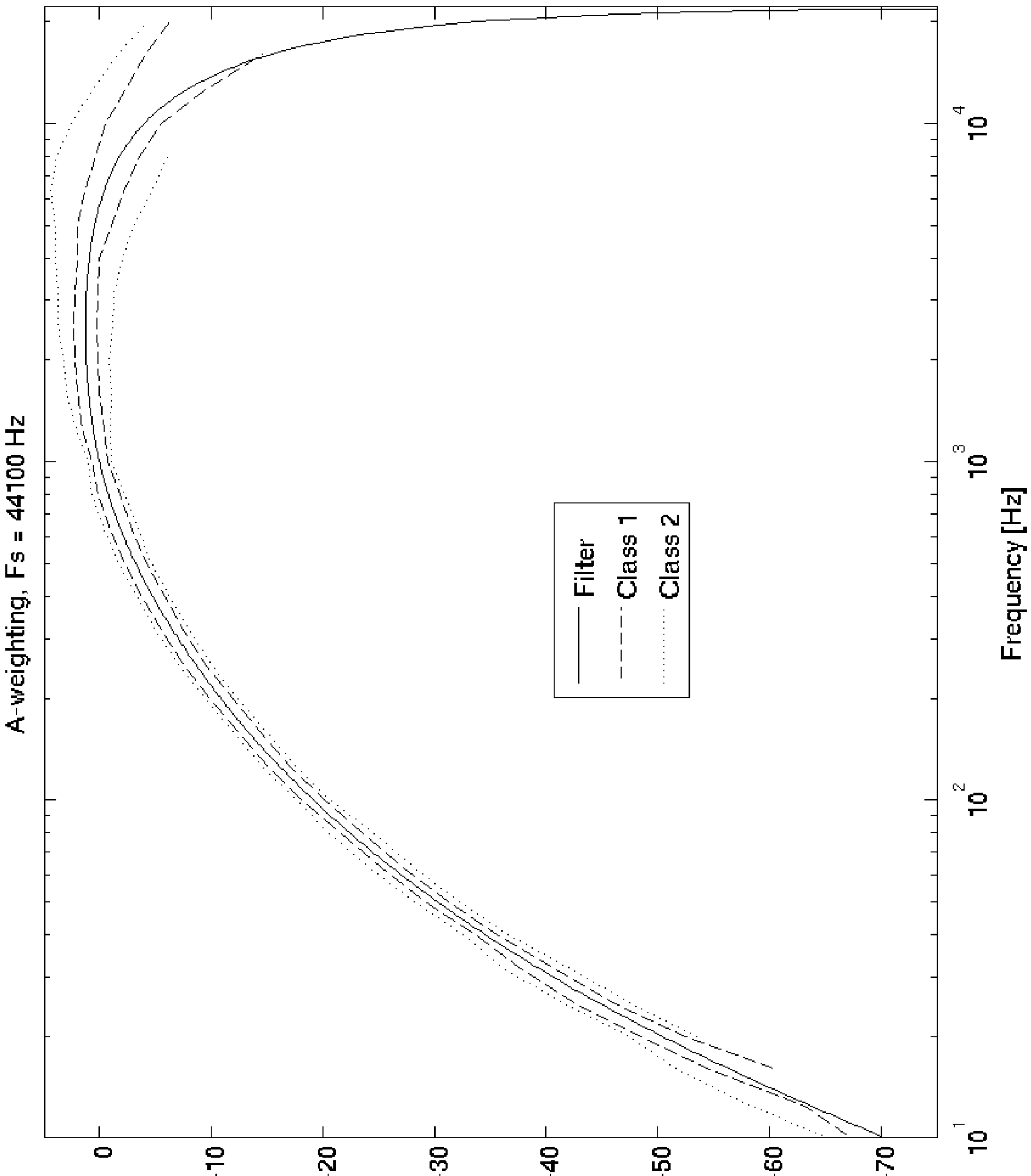

A-weighting, Fs = 44100 Hz
Filter
Class 1
Class 2
Gain [dB]
0
-10
-20
-30
-40
-50
-60
-70
Frequency [Hz]
10^1
10^2
10^3
10^4

FIGURE 10

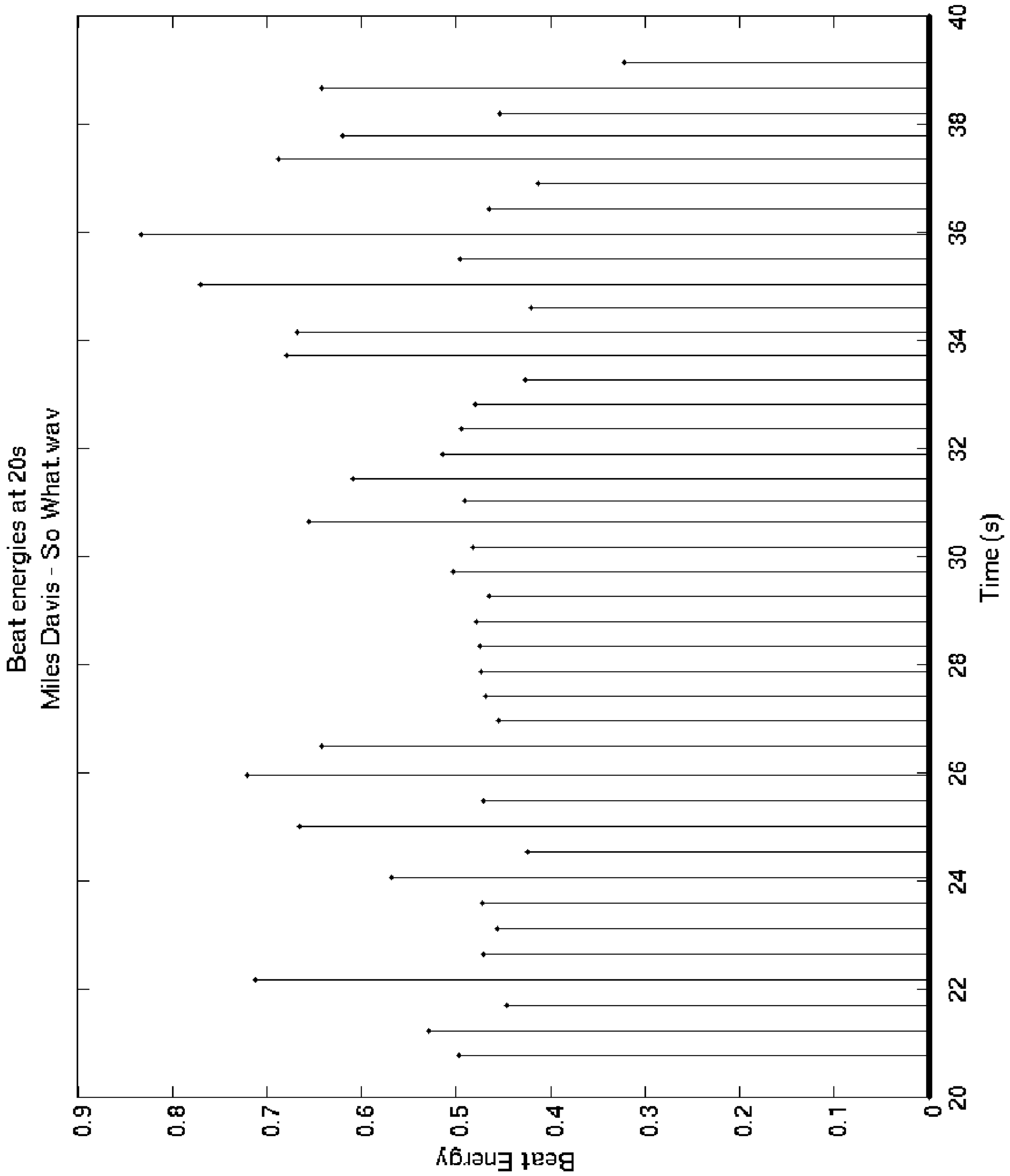

Beat energies at 20s
Miles Davis - So What.wav
Beat Energy
Time (s)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
20
22
24
26
28
30
32
34
36
38
40

FIGURE 11

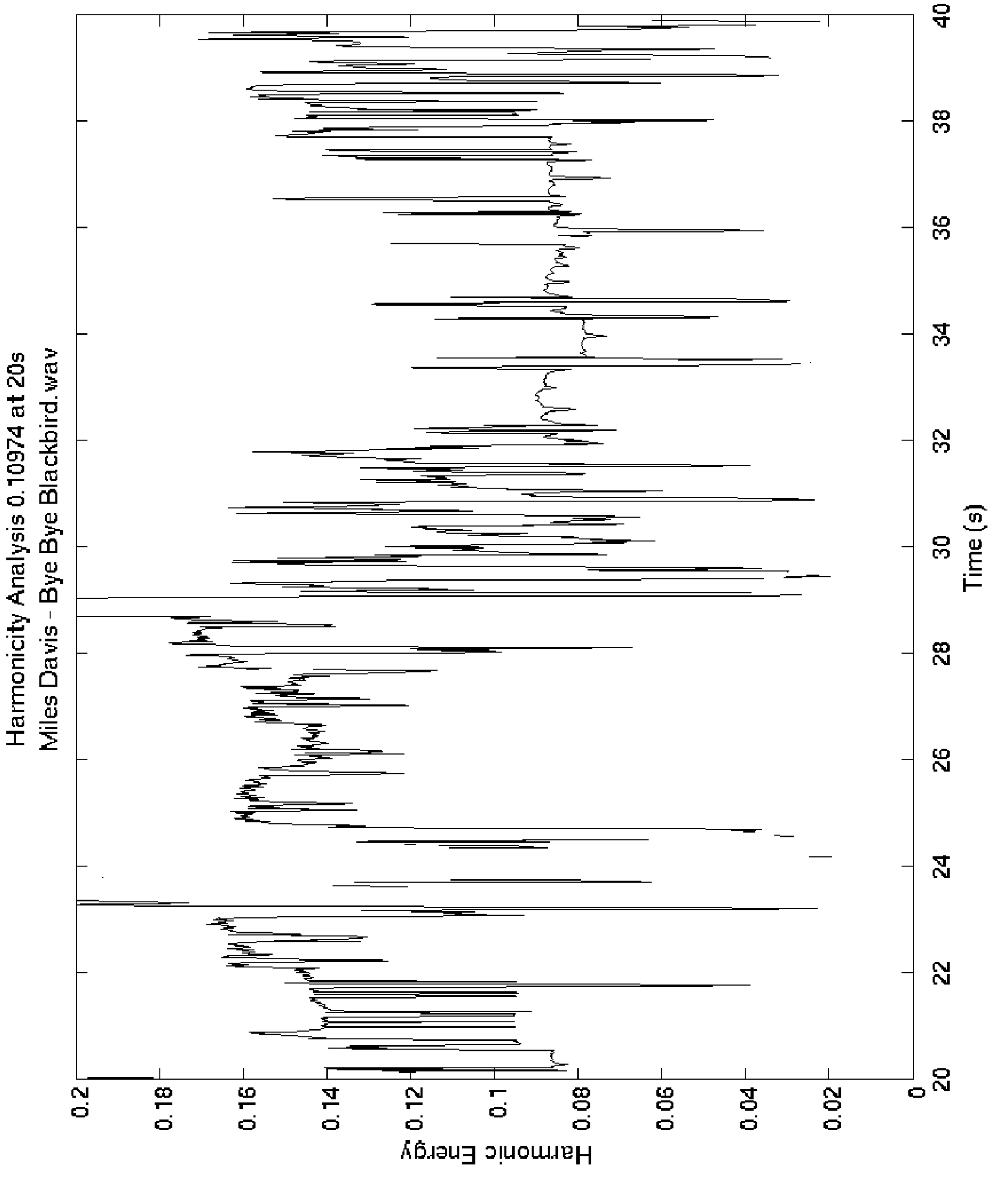

Harmonicity Analysis 0.10974 at 20s
Miles Davis - Bye Bye Blackbird.wav
Harmonic Energy
Time (s)
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
20
22
24
26
28
30
32
34
36
38
40

FIGURE 12

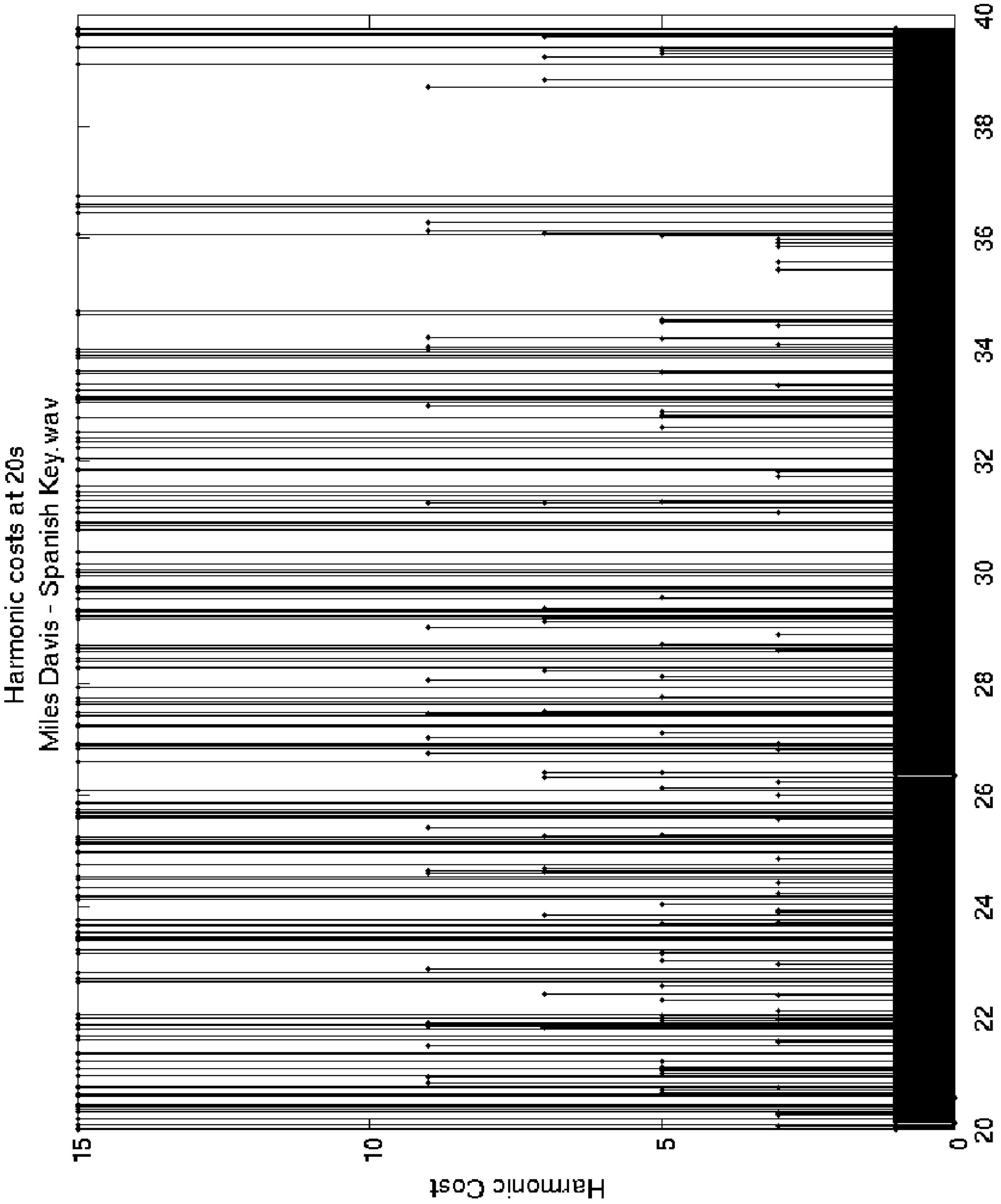

Harmonic costs at 20s
Miles Davis - Spanish Key.wav
Harmonic Cost
Time (s)
15
10
5
0
20
22
24
26
28
30
32
34
36
38
40

FIGURE 13

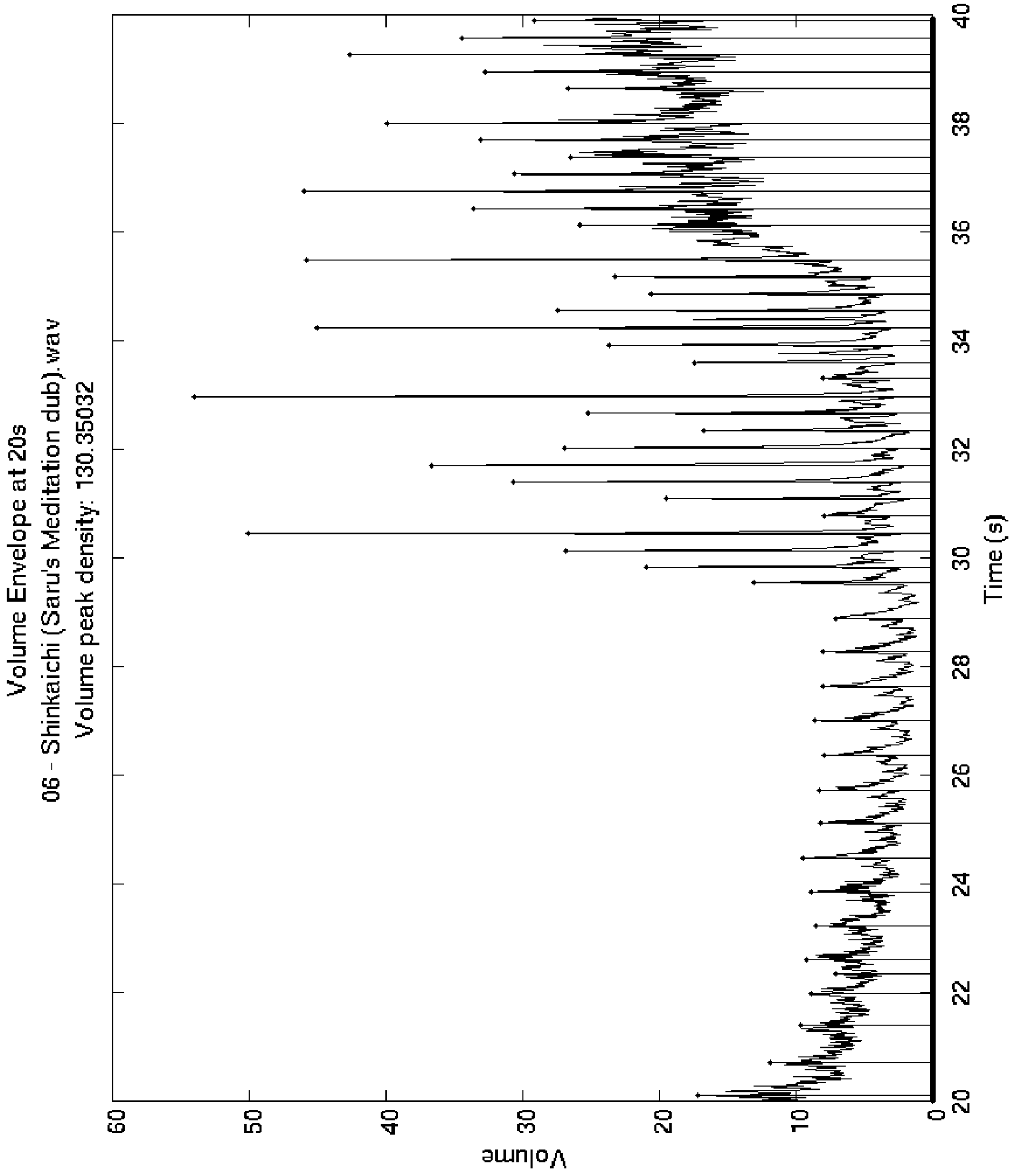

Volume Envelope at 20s
06 - Shinkaichi (Saru's Meditation dub).wav
Volume peak density: 130.35032
Volume
0
10
20
30
40
50
60
Time (s)
20
22
24
26
28
30
32
34
36
38
40

FIGURE 14

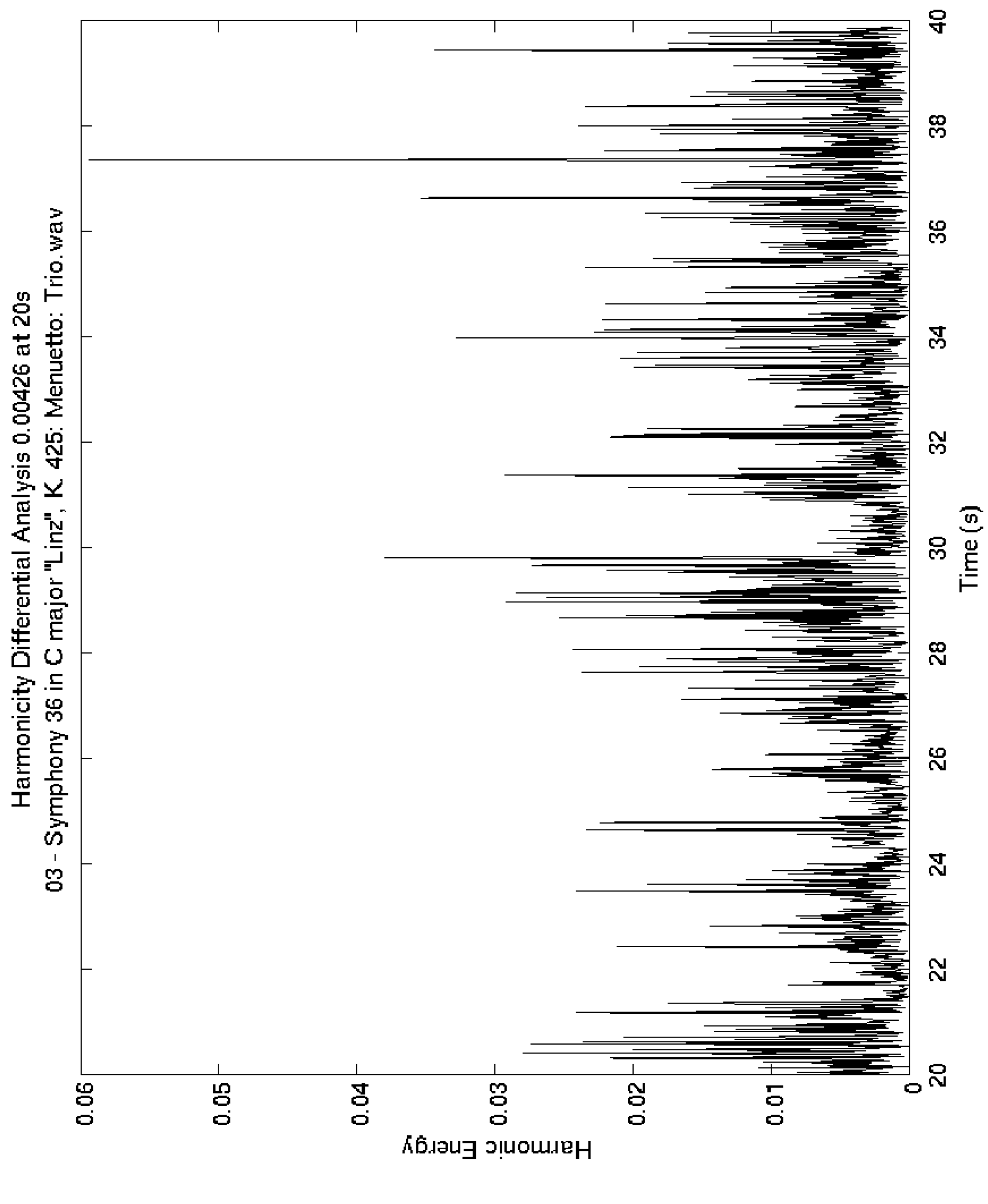

Harmonicity Differential Analysis 0.00426 at 20s
03 - Symphony 36 in C major "Linz", K. 425: Menuetto: Trio.wav
Harmonic Energy
0.06
0.05
0.04
0.03
0.02
0.01
0
20
22
24
26
28
30
32
34
36
38
40
Time (s)

FIGURE 15

Category 5 – high activation

Spanish Key	**Bitches Brew disc 2**	**duration 17.32**
Oleo	Half Nelson CD9	duration 5.52

Category 4 – moderately high activation

Bye Bye Blackbird	**Bye Bye Blackbird CD8**	**duration 7.55**
Surrey with the Fringe on Top	Four CD7	duration 9.06

Category 3 – normal (to high) activation

So What	**Kind of Blue**	**duration 9.22**
Ahmad's Blues	Half Nelson CD9	duration 7.27

Category 2 – moderately low activation

When I Fall in Love	**Half Nelson CD9**	**duration 4.25**
My Funny Valentine	My Funny Valentine CD10	duration 5.59

Category 1 – low activation

Flamenco Sketches	**Kind of Blue**	**duration 9.26**
Blue in Green	Kind of Blue	duration 5.37

FIGURE 16

Category	Rising	stable	falling
5.		•	•
4.	•	•	•
3.	•	•	•
2.	•	•	•
1.	•	•	

FIGURE 17

Category	Rising	stable	falling
5.		Symphony 7 in A, Op. 92 Allegro con brio	Symphony 7 in A major Op 92 Presto – Assai Meno Presto
4.	Symphony No 5 in C minor, Op. 67 Allegro con brio		Symphony 8 in F, Op. 93 Tempo di Menuetto
3.	Symphony 9 in D minor 'Ode to Joy', Op. 125 Allegro ma non Troppo, un Poco Maestoso	Symphony 6 in F major 'Pastoral', Op. 68 Awakening of cheerful feelings upon arrival in the country	Symphony 7 in A major, Op 92 Allegretto
2.	Symphony 3 in E flat 'Eroica', Op. 55 Allegro con brio		Symphony 6 in F 'Pastoral', Op. 68 Sheperds' song; cheerful and thankful feelings after the storm
1.	Symphony 3 in E flat 'Eroica', Op. 55 Marcia funebre: Adagio assai	Symphony 6 in F major 'Pastoral', Op 68 Scene at the Brook	

FIGURE 18

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2012/051314

A. CLASSIFICATION OF SUBJECT MATTER
INV. G10H1/00 G06F17/30
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G10H G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2006/050512 A2 (PLAIN SIGHT SYSTEMS INC [US]; COPPI ANDREAS C [US]; COIFMAN RONALD R [) 11 May 2006 (2006-05-11) abstract; figure 2 paragraphs [0011], [0012], [0043], [0044], [0052] - [0057] -----	54,65
X	US 2004/237759 A1 (BILL DAVID S [US]) 2 December 2004 (2004-12-02) abstract; figures 2A,2C,2D,3,4,5 paragraphs [0039], [0040], [0043], [0045] - [0047], [0052], [0061] - [0065], [0082], [0087] -----	54,65

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

25 September 2012

Date of mailing of the international search report

09/10/2012

Name and mailing address of the ISA/
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Feron, Marc

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2012/051314

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2007/131096 A1 (LU LIE [CN] ET AL) 14 June 2007 (2007-06-14) abstract; figures 2-5 paragraphs [0003], [0030] - [0033], [0038], [0051], [0059], [0060] -----	54,65
X	TOLOS M ET AL: "Mood-based navigation through large collections of musical data", CONSUMER COMMUNICATIONS AND NETWORKING CONFERENCE, 2005. CCNC. 2005 SECOND IEEE, IEEE, PISCATAWAY, NJ, USA, 3 January 2005 (2005-01-03), pages 71-75, XP010787613, DOI: 10.1109/CCNC.2005.1405146 ISBN: 978-0-7803-8784-3 abstract Chapter "III. Mood based retrieval -----	54,65
X	EP 1 703 491 A1 (SONY DEUTSCHLAND GMBH [DE]) 20 September 2006 (2006-09-20) abstract; figures 1A,1B,2,3 paragraphs [0029], [0035], [0044], [0048] -----	54,65
T	SHIHAB SHAMMA: "On the role of space and time in auditory processing", TRENDS IN COGNITIVE SCIENCES, vol. 5, no. 8, 1 August 2001 (2001-08-01), pages 340-348, XP55038877, ISSN: 1364-6613, DOI: 10.1016/S1364-6613(00)01704-6 the whole document -----	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB2012/051314

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 1-53, 55-64, 66-81
because they relate to subject matter not required to be searched by this Authority, namely:

see FURTHER INFORMATION sheet PCT/ISA/210

2. [X] Claims Nos.: 1-51, 53, 55-62, 67, 69-81
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

[ ] The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.1

Claims Nos.: 1-53, 55-64, 66-81

1. According to Art.17(2)(a)(i), Rules 39 and 67 PCT, this Authority is not required to examine claims 1-53, 55-64, and 66-81, the sole purpose of which is to perform mental acts, such as stimulating arousal, or to merely present information such as analysis results, or to claim diagnostic, therapy or business methods, or to claim subject-matter whose sole purpose is to provide a non-technical effect. 1.1. The subject-matter of claim 1 and all claims depending on claim 1, such as claims 2-51, 53, 55-62, 67, 71 and 76-81 is excluded from search under Rule 39 and 67 due to its scope being defined in claim 1 by the following result to be achieved: "so that appropriate sounds can be selected and played to a listener in order to stimulate and/or manipulate neuro-physiological arousal in that listener". This defines the problem to be solved as " stimulate and/or manipulate neuro-physiological arousal in that listener" by means of "appropriate sound selection" according to some scientific theory. This subject-matter of claim 1 is therefore excluded from search for each and any of the following three reasons (Rule 39.1 (iii,iv,v) PCT): 1.1.1. The subject-matter of claim 1 provides no technical effect, because its sole output is a mere presentation of information (Rule 39.1 (v) PCT), i.e. a presentation of the results of the analysis, as mere data, in an unspecified and unclear form, for non-technical future use. 1.1.2. The sole declared future use of that data aims at nothing but to trigger a specific mental act ("arousal") in a specific person:"that listener" (mental acts, see Rule 39.1 (iii) PCT; the fact that the result to be achieved is person-specific means that the claimed system has to be tailor made for that person and is therefore contrary to Art.33(4)PCT because it is not applicable in industry), 1.1.3. The result to be achieved shows that the category of the claim is incorrect: the claimed analysis system is defined solely ("analyses sounds so that...") by the therapy effect ("to stimulate and/or manipulate arousal")which is desired to be achieved; in other words it is an inextricable component of a therapy method and cannot be dissociated from that therapy method. However, such therapy method claims (methods of treatment of the human or animal body by surgery or therapy, see Rule 39.1 (iv) PCT) are excluded from search and are anyhow not applicable in industry. 1.2. Independent claims 52 and 63 give rise to the same objections for the same reasons, because like claim 1, they are directed solely to data analysis, i.e implicitly to mere a presentation of information about the analysis results, and therefore do not give rise to any technical effect. 1.3. Moreover the claims which depend on claims 52 and 63 and which also in and of themselves do not define or contribute a technical effect, such as e.g. claims 55-62, 64 and 76-81 are also excluded from the search based on the application of Rules 39 and 67 PCT. 1.4. Independent claim 66 gives rise to the same objections as claim 1 and is similarly excluded from search because "a method of categorizing sounds" yields a mere presentation of information, the only claimed application of which is clearly a therapy method: "to entrain neuro-physiological arousal towards a target level". 1.5. Independent Claim 68, because it claims "an automated diagnosis" is excluded from search as a diagnostic method for measuring arousal on the human or

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

animal body (Rule 39.1 (iv) PCT) 1.6. Independent claims 69 and 70 give rise to the same objections as claim 1 and are similarly excluded from search because "a computer implemented method of creating a playlist of tracks" based on "manually analysing musical parameters" yields a mere presentation of information based on performing mental acts (that's what "manually analysing" boils down to, a mere thought process) Another reason to exclude it from search is that its only claimed application and purpose of selection is clearly a therapy method defined as follows: "in order to entrain arousal and direct state of mind and/or affect" . 1.7. Independent claims 72-74 are claiming methods of recommending, selecting, categorising music based on vague, apparently non-technical criteria: they all correspond to mere presentations of information and mental acts (the use of a machine or device of any kind is apparently not necessary, these methods could be implemented by mere thought processes) which do not appear to be based on any technical effect, nor do they appear to provide any technical effect and therefore they are excluded from search for the same reasons as claim 1. 1.8 Independent claim 75: "a method of ordering a series of pieces of music" is excluded from search because it is in effect claiming a business method (Rule 39.1 (iii) PCT). Moreover, that ordering method could be implemented as a mental act, as it does not require any specific hardware.

Continuation of Box II.2

Claims Nos.: 1-51, 53, 55-62, 67, 69-81

A. The set of claims as a whole is contrary to the clarity, support and conciseness requirements of Art.6 PCT because of the excessive and unclear multitude of claimed inventive concepts and because none of the multitudinous inventive concepts in the set of claims as a whole is sufficiently clearly and concisely defined, nor areany of those concepts fully understandable from the disclosure as originally filed. Specifically, the claimed subject matter in claims 1-81 includes 19 independent claims (independent claims 1,52,54,63,65,66,68, 69 and 71-81) of unrelated yet partly overlapping scopes, so that it is not clear which among these 19 independent claims is intended by the applicant to represent the main inventive concept, to the extent that such a common inventive concept can be defined at all or understood at all, and to the extent to which each alleged concept can be understood as claimed and disclosed. B. However, claims 1-53, 55-64, 66-81 are also all directed, to at least some extent, to subject matter which this authority is not obliged to search under Rule 39 and 67 and which therefore will not be searched. Thus there is only one concept remaining to be searched, that of claims 54 and 65. C. The excessively multitudinous character is also evidenced by the catch-all nature of the wording of many of these claims (see in particular the wording of claims 71-81, especially claims 79 and 80 "or to form part of the system defined in any of the above systems" defining an unreasonably large number of alternatives) as well by the unreasonably large number of alternatives in their dependent claims. D. For instance the alleged inventive concepts of technical nature which would be the basis for search are really only defined by the

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

multitudinous dependent claims of claim 52 (since claim 52 is about an analysis without technical effect), among which the very unclear, unsupported claims 53,55-62 and 76-81 equally claiming multitudinous concepts, as in claim 55 when dependent on claim 52, claiming a further multitude of alternative, extremely vague, undisclosed and unsupported results to be achieved: " analysed for the purpose of navigation,or discovery,or retrieval, or selection, or matching for a specific requirement [which?], or for playlist creation, or for entraining affect, or for entraining mood". This multitude of unclear and unsupported concepts defined by claim 55 cannot be considered by this authority as a clear basis for any kind of search. These kinds of overly general and blatantly unsupported concepts will therefore simply be ignored. E. It follows that the alleged inventive concepts claimed in claims 56-62 and 76-81 are equally unreasonably multitudinous, for instance because they all depend on the impossible interpretation of the concepts claimed by this unclear and multitudinous claim 55. In particular, the set of claims as a whole does not comply with rule 6.1(a) and 6.4(a). F. The subject-matter of claim 1 and all claims depending on claim 1, such as claims 2-51, 53, 55-62, 67, 71 and 76-81 cannot be searched because it is is contrary to the clarity requirements of Art.6 PCT due to its being defined in claim 1 by the following unclear result to be achieved: "so that appropriate sounds can be played to a listener in order to stimulate and/or manipulate neuro-physiological arousal in that listener", and due to its being defined by unclear and unusual musical parameters defined by mere reference to an unclearly specified, undisclosed and unsupported model of the human brain. G. The subject-matter of claims 69 and 70 cannot be searched because it is is contrary to the clarity requirements of Art.6 PCT due to its being defined in claim 69 by the following unclear result to be achieved: "in order to entrain arousal and direct state of mind and/or affect".

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guideline C-VI, 8.2), should the problems which led to the Article 17(2) declaration be overcome.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2012/051314

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2006050512 A2	11-05-2006	US 2006111621 A1	25-05-2006
		US 2007270667 A1	22-11-2007
		WO 2006050512 A2	11-05-2006
US 2004237759 A1	02-12-2004	NONE	
US 2007131096 A1	14-06-2007	NONE	
EP 1703491 A1	20-09-2006	CN 101142622 A	12-03-2008
		EP 1703491 A1	20-09-2006
		JP 2006276854 A	12-10-2006
		US 2009069914 A1	12-03-2009
		WO 2006097299 A1	21-09-2006

(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 103890838 A
(43)申请公布日 2014.06.25

(21)申请号 201280039163.X

(22)申请日 2012.06.11

(30)优先权数据
1109731.8 2011.06.10 GB

(85)PCT国际申请进入国家阶段日
2014.02.10

(86)PCT国际申请的申请数据
PCT/GB2012/051314 2012.06.11

(87)PCT国际申请的公布数据
WO2012/168740 EN 2012.12.13

(71)申请人 X- 系统有限公司
地址 英国萨里郡

(72)发明人 尼格尔·奥斯本
罗伯特·阿什克罗夫特
保罗·罗伯森 彼得·金斯利

(74)专利代理机构 北京连和连知识产权代理有限公司 11278
代理人 李钦鹏

(51)Int.Cl.
G10H 1/00(2006.01)
G06F 17/30(2006.01)

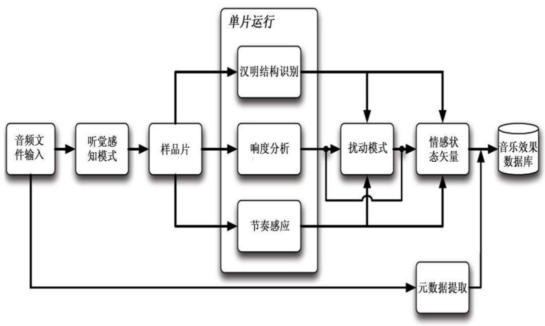
权利要求书7页 说明书36页 附图15页

(54)发明名称

用于分析声音的方法和系统

(57)摘要

本发明涉及分析音轨（例如音乐）的方法和系统。描述了人类大脑中的下层皮质、边缘和皮质下区域中的一个或多个的神经生理运行和对声音的反应的预测模型。声音被分析，以便合适的声音能够被选择和播放给收听者，以便刺激和/或操纵收听者的神经生理唤醒。该方法和系统特别适用于利用生物反馈资源的应用。



CN 103890838 A

1. 一种用于分析声音的计算机实现的系统，例如用于分析音轨，或任何其它类型的声音，其特征在于，该系统包括处理器，其编程为根据来源于或关联到人类大脑中的下部皮质、边缘和皮质下区域中的一个或多个的神经生理运行和对声音的反应的预测模型的音乐参数自动地分析声音；

并且其中，该系统分析声音，以便可以选择和播放合适的声音给收听者，以便刺激和/或操控该收听者的神经生理唤醒。

2. 根据权利要求 1 所述的系统，其特征在于，该系统适于自动分析声音并将该分析的结果存储在数据库中，以便随后能够从该数据库中选择合适的声音，并将其播放给收听者，以向该收听者提供神经生理唤醒的期望的刺激和/或操控。

3. 根据权利要求 1 或 2 所述的系统，其特征在于，音乐参数涉及节奏性。

4. 根据上述任意一项权利要求所述的系统，其特征在于，音乐参数涉及谐度，其是与谐波列的对应程度。

5. 根据上述任意一项权利要求所述的系统，其特征在于，音乐参数涉及扰动，其是音乐体验中的变化率和变化程度的测量。

6. 根据权利要求 3 和上述任意一项根据权利要求 3 的权利要求所述的系统，其特征在于，通过分析节拍感应，本系统利用具体校准的初始窗口预测性地模拟原始脊髓通路和运动前回路（例如，基底神经节、前庭系统、小脑），其所有都涉及对有节奏的脉动的原始反应。

7. 根据权利要求 3 和上述任意一项根据权利要求 3 的权利要求所述的系统，其特征在于，通过利用自相似/自相关算法，本系统预测性地模拟节奏模式识别和保持区域（例如，颞叶的次级听觉皮层）。

8. 根据权利要求 3 和上述任意一项根据权利要求 3 的权利要求所述的系统，其特征在于，本系统预测性地模拟镜像神经元系统的活动，其通过以下一个或多个检测节奏性活动的功率、轨迹和意向性：节奏功率的索引，包括音量级别的计算、音量峰值密度、“低谷”、或能量的缺乏和、运行能量的动态曲线。

9. 根据权利要求 3 和上述任意一项根据权利要求 3 的权利要求所述的系统，其特征在于，通过分析起始之前和起始中的能量消耗曲线（陡峭对应高唤醒，平滑对应低唤醒）、重要的镜像神经元信息，以及导致明显发音的能量流的曲线的计算，本系统预测性地模拟镜像神经元的活动。

10. 根据权利要求 4 和上述任意一项根据权利要求 4 的权利要求所述的系统，其特征在于，通过确定谐度和失谐度的水平，本系统预测性地模拟黑索氏回对声音的运行和反应。

11. 根据权利要求 4 和上述任意一项根据权利要求 4 的权利要求所述的系统，其特征在于，其通过谐波产物频谱的计算来检测主要的基音，之后建立不同基音频谱内和之间的谐度的等级。

12. 根据权利要求 11 所述的系统，其特征在于，将主要的基音的检测和谐度等级的确定“垂直地”应用于瞬时的时刻，和“水平地”应用于音高和频谱随时间的演进（涉及黑索氏回周围区域的音质映射），并且以线性谐波损失的方式表示，该谐波损失代表基音变化的速率和变化的谐波距离。

13. 根据权利要求 12 所述的系统，其特征在于，通过分析瞬时时刻的垂直谐度的窗口，

本系统预测性地模拟黑索氏回、颞上沟、岛环状沟对简单的音色的神经生理感觉。

14. 根据权利要求 12 所述的系统，其特征在于，本系统根据每个短时傅里叶变换（STFT）时间片偏离于谐波列的简单比例的程度来预测性地模拟旋律和谐波的演进：线性谐波损失从基音频率不同于之前的片的短时傅里叶变换（STFT）时间片处上升；在基音上无变化的时间片具有零损失。

15. 根据权利要求 5 和上述任意一项根据权利要求 5 的权利要求所述的系统，其特征在于，扰动将节奏性和谐度的变化索引与听觉脑干和皮质活动相结合，该听觉脑干和皮质活动神经支配杏仁核、海马和核心情感系统区域，该杏仁核、海马和核心情感系统区域影响神经传递和内分泌系统，包括肾上腺轴、多巴胺回路和例如去甲肾上腺素、褪黑素和催产素的水平。

16. 根据上述任意一项权利要求所述的系统，其特征在于，声音的分析在本地存储的音乐数据上实时地运行，并且该系统包括在个人计算装置上运行的软件、固件和 / 或硬件。

17. 根据上述任意一项权利要求所述的系统，其特征在于，声音的分析在存储在服务器中的音乐数据上实时地运行，并且该系统包括在该服务器或相关的服务器上运行的软件、固件和 / 或硬件。

18. 根据上述任意一项权利要求所述的系统，其特征在于，将定义针对特定的声音的节奏性、谐度和扰动的数据以给出代表唤醒的单一输出的方式结合，该声音例如为，音乐音轨或该音轨的部分。

19. 根据权利要求 3 和上述任意一项根据权利要求 3 的权利要求所述的系统，其特征在于，利用等式确定节奏性，该等式将 R 与 B 和 S 关联，例如等式 R = √ B*S^2，并且其中 R 是节奏性，B 是每分钟的节拍，S 是节拍强度的平均值。

20. 根据权利要求 3 和上述任意一项根据权利要求 3 的权利要求所述的系统，其特征在于，利用等式确定节奏性，该等式将 R 与 B 和 S 关联，例如等式 I = C/10 - H，并且其中 I 是失谐度，C 是线性谐度损失，和 H 是瞬时的节奏。

21. 根据权利要求 5 和上述任意一项根据权利要求 5 的权利要求所述的系统，其特征在于，利用等式确定扰动，该等式将 T 与 H 和 P 联系起来，例如 T = dH/dt*P，其中，T 是扰动，H 是节奏性，和 P 是峰值音量期间的能量。

22. 根据上述任意一项权利要求所述的系统，其特征在于，将针对给定的音轨的节奏性、谐度和——如果适用——扰动的值结合在一起，并映射到表征生理状态 E 的 n 维点 p 中。

23. 根据权利要求 22 所述的系统，其特征在于，该系统利用等式确定兴奋度，该等式将 E 与 I、R 和 T 联系起来，例如等式 E = (10*I*R)+T，其中，E 是兴奋状态，R 是节奏性、I 是失谐度，和 T 是扰动。

24. 根据权利要求 22 所述的系统，其特征在于，将针对给定的音轨的节奏性、谐度和——如果适用——扰动的值的变差映射到表征对生理状态的定向影响的 n 维点 p 中。

25. 根据上述任意一项权利要求所述的系统，其特征在于，分析声音是为了导航、或发现、或检索、或选择、或匹配特定需求、或创建播放列表、或诱导情感或诱导情绪的目的。

26. 根据上述任意一项权利要求所述的系统，其特征在于，声音是任意一种：音乐；工作间、影院、商店、车辆、汽车、火车、飞机的周围或背景音乐；自然的声音（风、海等），动物的声音、超声（行星、恒星、花、树、金融市场、细胞活动等）。

27. 根据上述任意一项权利要求所述的系统，其特征在于，该系统用于根据人类受检者预先选择的期望的唤醒状态来选择音轨并将其回放给人类受检者，根据人类对音乐的神经生理反应的模型来选择音轨，根据由神经生理模型预测的对音乐音轨的神经生理反应，神经生理模型用于选择回放的音乐音轨。

28. 根据上述任意一项权利要求所述的系统，其特征在于，该系统包含计算机，其可操作为根据预先选择的人类受检者期望的唤醒状态选择用于回放给人类受检者的音乐音轨，音乐音轨的选择是根据人类对音乐的神经生理反应的模型实现的，该模型在计算机上运行，神经生理模型由计算机使用，以根据由神经生理模型预测的对音乐音轨的神经生理反应，选择回放的音乐音轨。

29. 根据上述任意一项权利要求所述的系统，其特征在于，计算机可操作为接收人类受检者期望的唤醒状态的选择。

30. 根据上述任意一项权利要求所述的系统，其特征在于，人类受检者既选择他们期望的唤醒状态也是由系统产生的音乐音轨输出的接受者。

31. 根据上述权利要求 27 － 30 中的任意一项所述的系统，其特征在于，期望的唤醒状态的选择由不是人类受检者的一方执行。

32. 根据上述任意一项权利要求所述的系统，其特征在于，向用户呈现用户界面，这样，用户界面可操作为使用户能够从活动菜单中挑选，以便系统建立唤醒和情感的目标水平，这将促进选择活动。

33. 根据上述任意一项权利要求所述的系统，其特征在于，主要的音乐库和分析软件存在于人类受检者远程或本地地可操作的计算机上，其具有将音乐的选择传递给个人音乐播放器装置的能力，之后，该个人音乐播放器装置根据可用的音乐产生动态播放列表。

34. 根据上述任意一项权利要求所述的系统，其特征在于，通过机器学习完善对声音的人类神经生理反应模型，例如通过线性回归和 / 或神经网络方法。

35. 根据上述任意一项权利要求所述的系统，其特征在于，通过机器学习程序和统计分析完善对音轨的人类神经生理反应模型。

36. 根据上述任意一项权利要求所述的系统，其特征在于，人类受检者的生物指标数据利用传感器测量，以确定人类受检者的神经生理唤醒。

37. 根据上述任意一项权利要求所述的系统，其特征在于，该系统包含包括传感器的生物反馈系统。

38. 根据权利要求 37 所述的系统，其特征在于，生物反馈系统包含传感器、处理器和存储音乐分类数据的数据存储库，传感器设置用于测量人类受检者的一个或多个参数，并将所述测量传送给处理器，其中，处理器设置用于根据测量和音乐分类数据选择音乐片段。

39. 根据上述任意一项权利要求所述的系统，其特征在于，传感器用于测量人类受检者的唤醒状态，并且通过预测性模拟人类神经生理反应模型来分类的声音被流媒体式提供或以其它方式提供，以便实现人类受检者预先选择的期望的唤醒状态。

40. 根据上述任意一项权利要求所述的系统，其特征在于，该系统包含传感器，这样，一旦激活传感器，系统测量人类受检者的初始神经生理唤醒水平，并且该系统将自动地创建播放列表，其将首先反映出该唤醒水平，之后引导人类受检者接近并使他们保持在人类受检者的预先选择的期望唤醒状态。

41. 根据权利要求 36-40 中的任意一项所述的系统，其特征在于，传感器是以腕带，或改造的耳机，或脑电图（EEG）‘帽子’，或手套的形式。

42. 根据权利要求 36-41 中的任意一项所述的系统，其特征在于，传感器捕获包含生物计量参数的数据，其包括以下至少一个：心率，包括脉搏节律分析的心率、血压、肾上腺素和催产素水平、肌张力、脑电波和皮肤电导性。

43. 根据权利要求 36-42 中的任意一项所述的系统，其特征在于，系统包括传感器和软件包。

44. 根据权利要求 1 － 35 中的任意一项所述的系统，其特征在于，该系统包含软件但没有传感器，该系统依赖于平均预期的反应。

45. 根据上述任意一项权利要求所述的系统，其特征在于，创建播放列表，以便诱导或保持唤醒，以及引导精神状态和 / 或情感。

46. 根据上述任意一项权利要求所述的系统，其特征在于，个体的神经生理唤醒水平的测量被自动地进行并表示为数值，以便使其能够通过数据库中的理论上无限数量的音乐片段的任意一个的音乐效果所反映。

47. 根据上述任意一项权利要求所述的系统，其特征在于，该系统可操作为检测音轨的情绪唤醒参数信息，并且进一步可操作为将该信息嵌入到音轨或指向音轨的电子链接中，或作为与音轨关联的元数据。

48. 根据上述任意一项权利要求所述的系统，其特征在于，该系统可操作为能够实现对存储在远程或本地的数据库中的音乐的自动搜索，以查找具有满足定义的标准的标签的音乐。

49. 根据上述任意一项权利要求所述的系统，其特征在于，该系统可操作为在社交网络应用程序中分享唤醒值。

50. 根据上述任意一项权利要求所述的系统，其特征在于，该系统可操作为将与例如 Google®这样的搜索引擎交互的用户所关联的唤醒值分享给该搜索引擎，以便该搜索引擎能够在之后使用这些值来优化通过搜索引擎进行的搜索和 / 或广告选择。

51. 根据上述任意一项权利要求所述的系统，其特征在于，该系统可操作为将与浏览特定网站或网页的用户关联唤醒值分享给网站优化系统，以便网站优化系统能够使用这些值来优化网站和 / 或特定网页。

52. 一种分析音轨的方法，其特征在于，包含自动分析音轨的步骤，这是根据人类大脑中的下部皮质、边缘和皮质下区域中的一个或多个的神经生理运行和对声音的反应的预测模型。

53. 根据权利要求 52 所述的方法，其特征在于，包括根据需要的进一步的步骤，以使分析音轨的系统能够如上述权利要求 1 － 51 中任意一项所述的系统一样运行。

54. 一种根据人类受检者预先选择的期望唤醒状态来分析用于回放给人类受检者的音轨的方法，其特征在于，包含以下步骤：

(i) 存储一组可操作用于回放选择的单独的音轨；

(ii) 根据人类大脑中的下部皮质、边缘和皮质下区域中的一个或多个的运行和对声音的反应的神经生理模型，预测对单独的音轨的神经生理反应；

(iii) 接收人类受检者选择的期望的唤醒状态，和

(iv) 根据预测的对单独的音轨的神经生理反应和人类受检者选择的期望的唤醒状态，选择音轨。

55. 根据权利要求 52 － 54 中任意一项所述的方法，其特征在于，分析音轨是为了导航、或发现、或检索、或选择、或匹配特定需求、或创建播放列表、或诱导情感或诱导情绪的目的。

56. 根据权利要求 52 － 55 中任意一项所述的方法，其特征在于，将用户界面呈现给用户，进一步包含以下步骤：

(i) 该用户从用户界面中的活动菜单挑选

(ii) 该系统建立唤醒和情感的目标水平，其将促进所挑选的活动。

57. 根据权利要求 52-56 中任意一项所述的方法，其特征在于，进一步包含，根据音乐效果矩阵 M 中表示的值将音轨分类并将其编入索引的自动分类处理的步骤。

58. 根据权利要求 52-57 中任意一项所述的方法，其特征在于，进一步包含，针对它们的一般音乐的节奏性和失谐度的值来分析音轨的步骤。

59. 根据权利要求 52-58 中任意一项所述的方法，其特征在于，进一步包含，针对它们的一般音乐的扰动值来分析音轨的步骤。

60. 根据权利要求 52-59 中任意一项所述的方法，其特征在于，利用信号处理技术自动确定节奏性、失谐度和——如果适合——扰动的值。

61. 根据权利要求 52-60 中任意一项所述的方法，其特征在于，进一步包含，结合节奏性、失谐度和扰动的值以产生兴奋度或唤醒的测量的步骤。

62. 根据权利要求 61 所述的方法，其特征在于，兴奋度 E 等于 (10* 失谐度 I* 节奏性 R)+ 扰动 T。

63. 一种可操纵的计算机程序产品，其特征在于，当其在计算机上运行以分析用于回放给人类受检者的音轨时，该计算机程序产品应用基于人类大脑中的下层皮质、边缘和皮质下区域中的一个或多个的人类神经生理运行和对音轨的反应的预测模型的算法。

64. 根据权利要求 63 所述的计算机程序产品，其特征在于，人类神经生理反应的预测模型是人类神经生理反应的确定性模型。

65. 一种计算机程序产品，其特征在于，当其在计算机上运行以根据人类受检者预先选择的期望的唤醒状态来分析用于回放给人类受检者的音轨时，该计算机程序产品可操作为执行以下步骤：

(i) 识别一组可操作用于回放的选择的单独的音轨；

(ii) 根据神经生理模型，预测对单独的音轨的神经生理反应；

(iii) 接收人类受检者选择的期望的唤醒状态，和

(iv) 根据对单独的音轨的预测的神经生理反应和人类受检者选择的期望的唤醒状态，选择用于回放的音轨。

66. 一种将声音（例如不论流派和文化渊源的任何的音乐片段）分类的计算机实现的方法，其特征在于，该方法根据源于人类大脑中的下层皮质、边缘和皮质下的神经生理运行和对音乐片段的反应的预测模型的音乐参数，以这样一种方式实现，声音可以被选择（例如基于传感器捕获的生物计量数据自动进行）以诱导神经生理唤醒接近目标水平。

67. 根据权利要求 1 － 51 中任意一项所述的系统，其特征在于，其适用于提供在远程数

据库中的声音(例如音乐片段)的自动分类(例如,根据源于人类大脑中的下层皮质、边缘和皮质下的神经生理运行和对音乐片段的反应的预测模型的音乐参数)。

68. 一种个体的下层皮质、边缘和皮质下神经生理唤醒水平的自动诊断,其特征在于,将唤醒水平表示为一个值,以便与数据库中的理论上无限数量的音乐片段的任何一个的音乐效果相对应。

69. 一种计算机实现的方法,其特征在于,创建通过自动或手动分析音乐参数产生的音轨的播放列表,以便诱导唤醒和引导精神状态和/或情感,该音乐参数源于人类的下部皮质、边缘和皮质下的神经生理运行和对音乐片段的反应的预测模型。

70. 根据权利要求 69 所述的计算机实现的方法,其特征在于,包括以下进一步的步骤:

a)通过参考现有的描述性元数据,如果可用,例如,流派或用户创建的播放列表,挑选数据库中的音乐的子集;

b)通过将用户的下层皮质、边缘和皮质下神经生理唤醒的初始水平与包含在音乐效果矩阵的相关行中的音乐匹配,从该音乐子集选择一些片段,该片段将对应于用户的下层皮质、边缘和皮质下神经生理唤醒的初始水平;

c)选择目标的精神状态和/或情感;

d)选择一系列上升或下降的音乐效果值,其符合从初始到要求的神经生理唤醒水平的期望的诱导通路;

e)基于这一系列值,从音乐数据库中选择合格的内容;

f)随机地从合格的内容中以其它规则挑选播放列表,该规则例如,流派偏好、防重复规则或美国的数字千年版权法(DMCA)规则;

g)基于连续的或规则的或偶然的生物计量反馈,每隔一段时间重复计算播放列表。

71. 一种针对情感诱导的音乐(例如个人的)数据库的充足性的测定方法,其特征在于,利用权利要求 1 - 51 中任意一个所述的系统,并在之后显示关于充足性和不充足的信息给用户。

72. 一种推荐音乐内容的补充的方法,其特征在于,该方法针对个人音乐数据库,以便于保证数据库的充足性,该方法通过利用源于人类的下层皮质、边缘和皮质下的神经生理运行和对该音乐的反应的预测模型的音乐参数来实现。

73. 一种选择音乐的方法,其特征在于,其选择具有类似的音乐效果的音乐(例如根据源于人类的下层皮质、边缘和皮质下的神经生理运行和对音乐片段的反应的预测模型的音乐参数)。

74. 一种将音乐分类的方法,其特征在于,该方法根据其音乐效果而不是其描述性属性分类,利用源于人类的下层皮质、边缘和皮质下的神经生理运行和对音乐片段的反应的预测模型。

75. 一种给播放列表中的一系列音乐片段排序的方法,其特征在于,通过将每段的音乐效果与音乐效果矢量所描述的值的时间序列匹配实现该方法,该音乐效果矢量源于人类的下层皮质、边缘和皮质下的神经生理运行和对音乐片段的反应的预测模型。

76. 一种操纵用户唤醒的方法,其特征在于,通过利用上述任何一种方法和系统实现。

77. 一种在任何给定的环境中修改周围声音属性的方法,其特征在于,目的对收听者产生期望的神经生理反应,该方法通过利用上述任何一种方法和系统实现。

78. 一种适于执行上述任何一种方法的系统。

79. 软件、固件、系统芯片或音频堆栈，其特征在于，其编程为或适于执行上述任何一种方法或形成上述任何一种系统定义的系统的一部分。

80. 一种计算装置，例如智能手机或平板电脑，其特征在于，适于操纵用户的唤醒，该方法通过利用上述任何一种方法或通过利用或包括上述任何一种系统、软件、固件系统芯片或音频堆栈。

81. 适于与权利要求 80 所定义的计算装置一起工作的传感器。

用于分析声音的方法和系统

[0001] 发明背景

1. 技术领域

[0002] 本发明涉及一种分析声音的方法和系统(例如,音乐音轨)。来自于声音——例如音乐——数据库的音轨能够被分析,以便自动预测这些声音对收听者产生的效果和影响。

2. 技术背景

[0003] 公知地存在最适合例如学习、放松、睡觉或运动这样具体的活动特定神经生理唤醒水平(涉及情绪、精神状态和情感)。然而,因为这些唤醒水平是由意识心智、环境刺激、自主神经系统、内分泌活动、神经传递和基础代谢之间的复杂的相互作用引起的,所以其很难控制和维持。

[0004] 也公知地存在基于一系列复杂功能的对于音乐的普遍的人类反应,该一系列复杂功能从感知系统延伸,通过大脑皮层和其它的过程,以激活大脑和躯体系统的核心情感中心。同样公知这些功能存在于大脑的某部分中,例如,耳蜗、初级听皮层、运动前区皮层、杏仁核和导水管周围灰质(等)。例如节奏对运动前区皮层、自主神经系统、躯体系统、内分泌系统和神经传递具有可测量的影响。音乐结构和体验的其它方面也可影响人类的神经生理学,如下所述。

[0005] 3. 相关技术的论述

[0006] 已知三种对用于唤醒和逆唤醒人类的音乐的分析方法(为了简便起见,术语"唤醒"在本文中有时也包括逆唤醒)。第一种方法需要个人的判断,此人可以是专家或受检者本身。第二种方法是通过测试一些人并问其对于不同音乐音轨的感受。这两者都不可靠,因为每个都太过主观。

[0007] 第三种方法是分析根据音乐自身计算出的韵律(通常是节拍,但也包括平均能量的测量),并且将该韵律与所需的主体的唤醒状态关联。有几种这样的系统,其中一些将在下文中引用。其中大多数依赖于"诱导"(在惠斯登感应中,也就是与外部的节拍或节奏同步的倾向)或者增加的节拍(在已知的情况下,能量)和增加的作用或者唤醒之间的相关性(并且,对于减少的节拍和能量则相反)。

[0008] 现有技术的系统的实例是利用根据节拍选定的音乐来操纵唤醒和逆唤醒,包括专利文献 US282045、US191037、US113725、US270667、W0151116、US5267942。该技术可利用计算的每分钟的节拍数来预测诱导,或者,如专利文献 US060446 所述,可以调控节拍以便改进诱导。虽然该技术可以定向地校正,并且,通过惠斯登诱导原理(Huygens' entrainment principle)的扩展,其很可能通过一些曲目在一定程度上起效,但是,节拍难以被自动检测,并且其自身在容易而准确地检测节拍并且节拍接近收听者的当前心率(见下段)的受限情况下才可以最佳地用于计算神经生理影响。任何重要的分离和诱导的效果很可能会丢失。最重要地,如下所述,有效的节奏诱导不仅取决于每分钟的节拍数,并且不可分离地协同并依赖于其它的唤醒的音乐生成器,举个例子,例如谐度和扰动(turbulence)。

[0009] 专利文献US5667470是凭借对照曲目中确定的模式来实现或者否定音乐预期结果，而专利文献US4883067引入训练大脑的构思，以通过结合某些声音信号来复制神经活动的积极模式。一项专利文献US5267942，引用阿特舒勒在1948年记录的情绪同态原理(iso-moodic principle)作为其推断的证据，即，音乐节拍要对诱导心率产生任何影响，其必须位于个体的实际心率的“诱导范围”内，即接近它。其引入这样一个概念，即，一段音乐的神经生理影响依赖于受检者的初始状态，也就是说任何给定的音乐片段的影响是相对的而不是绝对的。也可以参考专利文献US2007/0270667，尝试使用生物计量反馈来操纵唤醒。

[0010] 也可以参考心理声学。心理声学已经广泛地应用在音乐压缩技术中（例如MP3），但另一个应用记录在专利文献US7081579中，其描述了一种基于七个测量的特性进行的歌曲相似性分析的方法，这七个特性是：亮度、频带宽、音量、节拍、节奏、低频噪声和八度音阶。在神经生理学术语中，这些技术能够识别“相似声音”的音乐（关于这点目前有许多），但是不能用于预测音乐的神经生理效果。

发明内容

[0011] 本发明是一种用于分析声音——例如音轨——的计算机实现的系统，该系统根据来源于或关联到神经生理运行的预测模型的音乐参数以及人类大脑中的下层皮质(lower cortical)、边缘(limbic)和皮质下区域(subcortical region)中的一个或多个对声音的反应来自动分析音乐；

[0012] 并且在其中，该系统分析声音，以便可以选择和播放合适的声音给收听者，从而刺激和/或操纵该收听者的神经生理唤醒。

[0013] 该模型是一种“人类神经生理运行和响应的预测模型”，因为其预测大脑（例如，在下层皮质、边缘和皮质下的区域中的结构，包括相关的自主神经系统、内分泌系统、和神经传递系统）将如何响应特定的声音。

[0014] 在一个实施例中，分析了来自于音乐数据库的音轨，以便自动预测这些声音将给收听者带来的神经生理效果或影响。由此能够选择不同的音轨及其最佳的播放顺序，以操纵神经生理唤醒、精神状态和/或情感——例如接近、达到或保持所需的唤醒或逆唤醒状态、精神状态或情感（术语“情感”用在情绪、心情或状态的心理感受中）。

[0015] 我们可以将该系统与传统的心理声学（是例如MPEG MP3音频压缩算法的基础）进行比较，因为心理声学总体上涉及通过模拟信号处理来了解如何处理引入的压力波，该信息的处理是由例如耳蜗和初级听觉皮层承担，然而，本发明涉及声音效果——例如大脑中的下层皮质、边缘和皮质下的区域的神经生理运行以及对声音的反应。并且，心理声学科学不涉及为了刺激和操纵所需的收听者的唤醒状态而选择特定声音。

[0016] 我们也可以将该系统与音乐效果的普通模型(trivial model)进行比较，例如增加的节拍导致更大的唤醒。从这样的模型中完全缺失的是对神经生理运行和对声音的反应的广义理解；而且，在实际中，这样的模型过于薄弱以至于不具备真正的预测特性，并且，基于上述原因，该模型不是针对选择不同声音以便刺激和操纵收听者的唤醒水平的技术问题的通用的解决方案，这与本发明不同。

[0017] 源于或关联到预测模型的音乐参数可涉及节奏性和谐度也可涉及扰动——这些术语将在下面详细解释。本发明可用于音乐的搜索、选择、定序（即按顺序安排）、使用、推

广、订购和销售。其可进一步用于选择、修改、定购或设计非音乐的声音，以对收听者产生期望的神经生理效果，或可以用于允许选择，例如设计或修改发动机排气音调、电影配乐、工业噪声和其它的音频源。

[0018] 本发明通过一种称为 X- 系统的系统实现。X- 系统包括音乐音轨的数据库，根据源于或关联到人类神经生理运行的预测模型的音乐参数以及对那些音轨的反应已经分析了该音乐音轨。X- 系统也可以包括传感器、用于选择合适的音轨的音乐选择算法 / 播放列表计算器以及连接音乐播放器的连接器。一旦激活传感器，系统将诊断受检者的神经生理唤醒的初始水平，并自动构建源自于对 X- 系统编码的音乐或声音数据库的搜索的播放列表，该播放列表将首先符合或反映此唤醒水平，之后将收听者引导向和帮助将他 / 她保持在期望的唤醒水平。按照需要，根据定期的神经生理或其它指示信号的测量而重新计算播放列表。

[0019] 神经生理状态的测量可以利用多种技术实现，例如电子脑造影、正电子放射断层造影术、血浆、唾液或其它的细胞采样、皮肤电导、心率以及一些其它的技术，同时反应的预测可以通过任何合适的算法集得以实现，其首先假设，之后通过测试来完善。任何给定的算法集将依赖于模拟的刺激以及测量刺激效果的生物计量，但是，即使给定常量参数，也有很多有效的数学方法 :因此在本说明书中所述的特定的算法本身不是本发明最根本的特征，即使该系统中的大部分算法在构思和实施上都是独特的。也没有选择特定的生物计量来测量神经生理状态，尽管皮肤电导和心率都适合于一般用途，这是因为它们使测量能够容易地和非侵入性地实施，同时能够给出对自主神经系统中的唤醒或逆唤醒的良好指示，这又在很大程度上协同于内分泌活动和相关的神经传递。

[0020] X- 系统代表了基于现有技术的改进，因为 :a) 通过参考大脑的音频刺激处理，该音频刺激包括音乐，描述了音乐的生物活性成分（除了节拍和能量以外），和 b) 描述了如何将任何给定的声源校准到受检者的初始状态，以便具有最大化的诱导效果。其提供了很多优于其它系统的优点，其既不需要节拍的调制（从专利文献 US2007/0113725、US20070060446A1、US2006/0107822A1 已知节拍调制）也不需要心理声学校正的成分、合成音乐（从专利文献 US4883067 已知）来实现其效果。在无需以任何方式操纵音乐的渲染的情况下，X- 系统提供了利用全世界的音乐曲目进行情感的调制的可能性。

[0021] X- 系统是以我们将其称之为“对音乐的先天神经生理反应”(INRM- 我们将在下面详细描述）的范例为基础的，并且以与这些反应相关的下层皮质、边缘和皮质下功能中的一个或多个的独特信息模拟为基础。X- 系统具有独特的能力来自动分析音轨并构建产生收听者的唤醒和逆唤醒水平的可能性。这一独特的分析方法是人类通用的分析方法，并且其不但可以适用于所有人类文化的音乐而且适用于环境的和其它的声源。X- 系统有能力根据核心情感效应将音乐和声音数据库归类。X- 系统可远程地实施自动分类，例如针对个人的曲目。利用独特的基于无线电电极和扩音器的电导 / 心率传感器和其它装置，X- 系统也可以具有检测用户的心理状态和身体的能力。X- 系统可使用该传感器数据从任何选择的曲目中子选择音乐，无论是通过单独的音轨还是诱导的序列，当听这些曲目时，将帮助用户实现兴奋、放松、专注、警觉、提高身体活动的潜力等目标状态。其实现是通过分析用户的音乐数据库中的音轨（利用源于人类神经生理反应的预测模型的音乐参数），并且之后自动创建音乐播放列表，其也可以根据实时的生物反馈被动态地重新计算，该音乐播放列表将播放

给用户，以便引导她／他接近，且有助于保持她／他处于期望的目标状态。

[0022] 如上所述，X-系统模拟音乐对大脑下部和中部的特定部分——包括边缘系统和皮质下系统——的效果，但是大脑对音乐做出反应的部分不仅仅是这些。其他的中心支配着更多的个人感受，包括偏好、修养、记忆和联想、歌词的意义、写下它们的历史背景、表演者或作曲家有关情形的认识以及其它的因素。这些也具有显著的效果，因此重要的是不要期待任何音乐片段对任何个体具有绝对的影响。INRM 描述了音乐效果的重要部分，但不是全部。某些音乐片段将使收听者平静、甚至诱发睡眠的预测不像药物或麻醉剂，其中某种剂量的效果可以以合理的精确性预测到，并且该效果不能被有意识的努力所抵抗。然而，实验证实了 INRM 模型所基于的大脑的每个元素紧密的联系着唤醒和逆唤醒。然而，当适当地选择以伴随所需的状态或活动时音乐具有其最大的效果，并且 X-系统提供了选择总是适合于收听者正在做的事情的音乐的自动化方法，这在许多情况下可以非常有效，从治疗焦虑症以增强松弛或专注，或刺激创造“流”，或为体育活动带来力量和流畅性。支持 X-系统的大脑模拟提供其它存在的分类系统所不能提供的进一步能力：就是普遍性；X-系统可针对全世界的音乐曲目准确地预测心理唤醒的水平，无论是西方古典和流行、中国或印度的古典或民间的音乐、非洲流行或传统、还是前卫的电子或爵士乐。

[0023] 已经证明，在广泛的曲目中，X-系统在根据心率和皮电反应电阻这样的生物计量参数预测唤醒／逆唤醒的总指数（general index）方面能够胜过音乐专家，但是在这些生物计量参数不同的情况下，等式——我们将在本文的后面部分中描述——几乎确定需要修改。同样地，有很多本领域技术人员熟悉的数学技术，其已经用于预测一段音乐的神经生理效果，并且许多中的任何一项可产生同样令人满意的结果。因此，本发明的关键特征在于对音乐中的模式的识别，其是有神经生理活性的（“生物活性”）且对于人类的神经生理——包括唤醒和逆唤醒——具有可预测的效果。

[0024] 本发明的其它方面

[0025] 我们下面列举了本发明 15 个进一步的方面，每个也可以与其它任何一个相结合：

[0026] 1. 一种以这样一种方式将声音分类的通过计算机实施的方法（例如无论流派或文化渊源的任何音乐片段）（例如根据源于人类的下层皮质、边缘和皮质下的神经生理运行和对音乐片段的反应的预测模型的音乐参数），即，其可以被选择（例如根据由传感器捕获的生物计量数据自动地进行）以用于诱导神经生理唤醒接近目标水平；这可能发生在引导收听者接近一些预先指定的精神状态和／或情感中的一个或多个的时候，或者为了引导收听者接近一些预先指定的精神状态和／或情感中的一个或多个而发生。

[0027] 2. 在远程数据库中（例如根据源于人类的下层皮质、边缘和皮质下的神经生理运行和对音乐片段的反应的预测模型的音乐参数）的声音（例如音乐片段）的自动分类。这包括我们能够搜索／发现具有类似于 X-系统深层结构的的音乐和将传统的分类方案（Gracenote®等）与 X-系统交叉匹配的计划。作为自动分类的另 - 种选择，或者除此之外，X-系统提供了用于商业和宣传目的的选择和“推送”，或者用于特定音乐的描述和检测的方法论，其针对所有的应用，而不仅是诱导。一个实例是一种根据对音乐的先天神经生理反应将不论流派或文化渊源的任何音乐片段分类的通过计算机实现的方法，以用于搜索、引导、音乐发现、检索和选择的目的。

[0028] 我们现在扩大搜索／发现的概念，其中 X-系统提供远程或者本地音乐数据库以及

X- 系统编码服务的自动搜索。在本发明中，用户可以 ：

[0029] ●通过在其计算机或者智能手机的 X- 系统设备应用程序上按下“查找更多”或“我喜欢”键，就可以搜索具有与他们标记为喜欢的音乐类似的签名的音乐。

[0030] ●通过社交网络群体中的收听喜好模式搜索和搜索社交网络群体中的收听喜好模式，这样通过分享自己的喜好和选择并且将这些与自己的朋友交流，人们将看到自己对特定音轨的情绪反应和在网络中与其它人的对照之间的关系。

[0031] ●通过音乐或体验之旅搜索，这样可以存储特定的音乐序列在例如自己的智能手机中，并且当自己按下“我喜欢该序列，存储其以便我再次播放”时重复。

[0032] ●通过发现模式和用户标记为“我喜欢”的音轨之间的关系搜索，这样所述流派、音乐家、活动的类似的组合和 X- 系统编码的唤醒数据能够主动进行推荐。因此，例如，X- 系统将产生播放列表建议，其将结合爵士、特定的迈尔斯·戴维斯音轨、记录一篇短文、专注和唤醒水平，其前提是类似的组合已经从更早的收听序列中标记（ 标记的活动是智能手机应用程序的一部分 ）；和

[0033] ●在 Google 和其它网站搜索 X- 系统编码的信息，这样，例如，对音乐、视频或其它网页内容进行分类和标记——这是自动地进行 ；或者与搜索引擎服务商合作，以便其“宣传”X- 系统唤醒或情绪状态 ；或者根据在查看网页时自动标记网站的访问者。

[0034] 3. 个体的下层皮质、边缘和皮质下神经生理唤醒的水平的自动诊断并将其表达为数值，以便对应于在数据库中理论上数量无上限的音乐片段中的任何一个的音乐效果。可选择地或附加地，可以提供一种自我诊断的反复试验的方法，例如通过上述的歌曲选择。

[0035] 4. 一种创建音轨播放列表的计算机实施的方法，该音轨播放列表是由自动（ 或实际上手动 ）分析音乐参数产生的，该音乐参数源于人类的下层皮质、边缘和皮质下的神经生理运行和对音乐片段的反应的预测模型，以便诱导唤醒和引导精神状态和 / 或情感。可选择地，其可包括 ：

[0036] 通过参考现有的描述性元数据——如果可用，例如流派或用户创建的播放列表——来在数据库中挑选音乐的子集 ；b) 通过将用户的下层皮质、边缘和皮质下的神经生理唤醒的初始水平与音乐效果矩阵的相关行中包含的音乐进行匹配（ 我们将在之后详细解释该矩阵 ），从该音乐子集中选择符合该神经生理唤醒的初始水平的一些片段 ；c) 选择目标的精神状态和 / 或情感 ；d) 选择一系列升序或降序的音乐效果值，其对应于从初始的到所需的神经生理唤醒水平的预期诱导通路 ；e) 在这一系列的值的基础上，从音乐数据库选择合格的目录 ；f) 从合格的内容中随机挑选播放列表，该合格的内容服从于其他的规则，例如流派喜好、防重复规则（ 看下面的“音乐选择算法”）或美国的千禧年数字版权法 (DMCA) 规则 ；g) 以连续的生物计量反馈为基础每隔一段时间重复播放列表的计算，例如，基于包括最新反馈的生物计量反馈，播放列表可以被每分钟重复计算一次。

[0037] 5. 一种确定音乐数据库（例如个人的）对于情感的诱导的充分性并且之后将关于充分和不充分的信息显示给用户的方法。

[0038] 6. 一种推荐音乐内容的补充给个人的音乐数据库以便保证充分性的方法，这是通过使用源于人类的下层皮质、边缘和皮质下的神经生理运行和对该音乐的反应的预测模型的音乐参数实现的。

[0039] 7. 一种选择具有类似的音乐效果的音乐的方法（ 例如根据源于人类的下层皮质、

边缘和皮质下的神经生理运行和对音乐片段的反应的预测模型的音乐参数）。其可包括X- 系统编码的搜索。

[0040] 8. 一种根据音乐效果而不是描述性特性的将音乐分类的方法。

[0041] 9. 一种通过将每段的音乐效果与音乐效果矢量所描述的值的时间序列匹配以便给播放列表中的一系列音乐片段排序的方法。

[0042] 10. 一种利用上述任何一种方法和系统操纵用户唤醒的方法。

[0043] 11. 一种利用上述任何一种方法和系统来修改任何给定的环境中的环境声属性以便在收听者中产生所需的神经生理反应的方法。以及其作为定义这些反应的选择、控制或设计的工具的用途。

[0044] 12. 一种适于执行上述任何一种方法的系统。

[0045] 13. 北边称为或者适应于执行上述的任何方法或形成上述系统的一部分的软件（不论是装置驻留、网络驻留或其它位置）、固件、系统芯片 (SoC) 或音频堆栈 (audio stack)。

[0046] 14. 一种适于通过利用上述任何方法或通过利用或包括上述任何系统、软件、固件、系统芯片或音频堆栈来操纵用户的唤醒的计算装置，例如智能手机或平板电脑。

[0047] 15. 适于与上述定义的计算装置一起工作的传感器。

[0048] 一些更广义的说明如下：

[0049] 关于音乐中的哪些结构和经验现象 (experiential phenomena) 激活原脑 (primitive brain) 的哪些部分的识别，利用数字签名分析来测量的技术的发展，以及一系列通用模型结构的构建，其使用相对简单的等式预测大脑的相关区域和器官的激活水平并且进而预测其在生物计量指标上效果，这些是本发明的关键方面。

[0050] 本发明的实例可对所有的音乐流派有效并且不依赖于任何预存在数字化音乐数据库中的元数据。该数据库可以由用户从他或她自己的收藏中汇编并且存储在本地播放器中，在这种情况下，数据库中的音乐可以进行远程地描述，其可以在数字存储装置上被提供预分析，或其可以从中央服务器流动式接收。在后一种情况下，音乐可以与其它的数据和 / 或数字媒体关联，以便增强用户的体验，或者可以描述和包括签名片段 (signature excerpt)，以便促进所需的效果。

[0051] 本发明可以作为应用软件在远程服务器、音乐播放器自身或另一个直接或通过本地或广域网连接音乐播放器的装置，或固件或嵌入在芯片中执行；其可以形成音频堆栈的一部分或用作一组设计工具的一部分。这些实施例可实现音轨和其它声音的实时分析，都在例如智能手机或平板电脑这样的便携式计算装置内部本地地完成，或在服务器上远程地完成，或一些分散式的基于本地和服务器的处理的组合。所有这些部署也将支持一致的 API（应用程序界面），使应用程序供应商和服务提供商有接入系统的能力，例如，使新的应用程序被构建和部署。

[0052] 如果必要的元数据是可用的，可在音乐数据库上的这些元数据中挑选首选的音乐风格；如果必要的元数据不是可用的，系统也可从全部音乐数据库而不是挑选的子集中选择。

[0053] 如下术语在本文中采用特定含义：

[0054] “神经生理唤醒的水平”：一种计算出的指数，其根据例如皮肤电导和脉搏率来计

算,虽然也可以选择其它的参数,包括需要更复杂的测量的情况。不同的神经生理唤醒水平促进不同的活动、精神状态和情感。

[0055] “精神状态”:与例如创造力、学习、沉思、想象等这样的不同类型的思维相关的大脑的功能区之间的动态关系。

[0056] “情感”(名称):在心理学中使用以表示感觉或情绪,以及在精神病学中表示表达出的或观察到的情绪反应。心情。

[0057] “音乐效果”:由给定的音乐片段引起的精神状态或心情,以及其对神经生理唤醒的影响。

[0058] “声音”:包括任何声音,其包括如传统地理解的音乐,但是也延伸到其它的声音,例如,在工作场所、电影院、家、商店、车辆、汽车、火车、飞机——声音可以在理论上影响收听者的唤醒的任何地方——的环境或背景噪声。例如,调谐汽车排气的音律就是一个实例;修改引擎声是另一个实例。自然的声音(风、海等)、动物的声音、超声(行星、恒星、花、树、金融市场、细胞活性等)是“声音”的其它实例。在本文中,我们将涉及“音乐”,但是该术语应该被宽泛地解释为,不仅包括艺术形式的音乐,其中的声音和/或乐器联合展现了和谐、优美或自我表现,而且还包括其它所有的声音形式,如上述宽泛定义的术语。

[0059] 关于术语的说明:初级听觉皮层位于新皮质——大脑中最“进化”的部分——的颞叶中,但是其实质上在系统的“下层”,因此是“下层皮质”。对X-系统关键的器官——例如海马和杏仁核(amygdala)——通常被描述为“边缘”(来自拉丁文“limen、liminis”,意思是“阈值”,即,在新皮质的下限部分)。这与情绪相关的区域接近,例如伏隔核(nucleus accumbens)、和导水管周围灰质有时也被视为边缘。边缘系统也可以描述为原脑皮质(archicortex)和旧皮质(paleocortex)——“主要的、最初的或支配的”和“旧的”皮质。最终,涉及节奏,核心情感和运动的很多X-系统区域是皮质下,例如基底神经节和小脑。

[0060] 因此,相对于新皮质的更多认知相关、文化相关和反射区域来说,X-系统主要涉及大脑的下层皮质、边缘和皮质下区域,关系到对音乐的基本的和普遍的反应。

附图说明

[0061] 图1表示涉及适用于X-系统的音频处理的神经元件的图解示意图。实线框内封闭的元件是现有模型的一部分;包含在虚线框中的元件可以包括在该模型中。

[0062] 图2表示其中系统的用户既要选择他们所需的情感目标又是产生的输出的接受者的整个系统结构。

[0063] 图3表示其中由系统用户以外的一方做出目标情感的选择的整个系统结构。

[0064] 图4表示X-系统发明的一个实施例,其中,软件的各个方面都位于用户的个人电脑上(术语“个人电脑”应该被宽泛地解释为涵盖任何形状因素的任何计算装置,包括能够执行计算功能的任何装置)。

[0065] 图5表示X-系统发明的一个实施例,其中,初级的音乐库和分析软件位于用户的个人电脑上,其有能力传递选择的音乐给个人音乐播放器,其之后基于可用的音乐产生动态播放列表。

[0066] 图6表示X-系统发明的一个实施例,其中,外部服务提供商通过网络连接提供分析工具。音频可位于用户的个人电脑上或由服务提供商“流动式接收”,并且存储音乐情感

的数据库可用于最小化音轨分析。

[0067] 图 7A 是表示 X- 系统音频分析工具的用于分析谐度的主要部分的详细的方框图。

[0068] 图 7B 是表示 X- 系统音频分析工具的所有主要部分的详细的方框图。

[0069] 图 8 是表示 X- 系统音乐重放和监测应用程序的主要部分的详细的方框图。

[0070] 图 9 表示以图表表示针对兴奋、平稳或放松的通路的随时间变化的唤醒。

[0071] 图 10 :通过使用 A- 加权滤波器实现耳蜗和初级听觉通路的模拟。其弱化低频和增大高频,向着人类听觉的频率上限再次快速地下降。

[0072] 图 11 表示随时间变化的节拍能量 (Beat Energy)。

[0073] 图 12 表示随时间变化的谐波能量 (Harmonic Energy)。

[0074] 图 13 表示随时间变化的谐波损失 (Harmonic Cost)。

[0075] 图 14 表示随时间变化的音量。

[0076] 图 15 表示随时间变化的谐波能量。

[0077] 图 16 表示从迈尔斯·戴维斯的曲目中的取样分类。

[0078] 图 17 表示其它手动分类的实例,其中,音轨被进一步分类成稳定、上升和下降的矢量。

[0079] 图 18 表示来自贝多芬交响曲的乐章中已经根据矢量分类的实施例。

具体实施方式

[0080] 具体实施方式由以下部分组成 :

[0081] A. 高级概念

[0082] B. 对音乐的先天神经生理反应 (INRM) 的详细解释

[0083] C. 如何使用 X- 系统

[0084] D. 一个或多个传感器

[0085] E. 音乐选择算法

[0086] F. 音乐播放器

[0087] G. 诊断和成流软件

[0088] H. 手动分类

[0089] I. 手动分类矢量

[0090] J. 社交网络

[0091] K. 扩展 / 增强的时机

[0092] L. X- 系统的优势。

[0093] A. 高级概念

[0094] 科学证明,通过直接神经生理的参与,音乐诱导和塑造唤醒、精神状态和情感 ;本发明关注对音乐的先天神经生理反应的测定的发现和一般方法,并且包括一种利用这种现象的新方法。如上所述,本发明通过称为 X- 系统的产品实现。X- 系统利用音乐的潜力来影响收听者的神经生理变化,尤其涉及唤醒和逆唤醒以及相关的精神状态,致力于大脑的边缘、下层皮质和皮质下区域的最基本的、先天的神经生理活动和反应的水平。

[0095] 其不同于其它的音乐分类方法,因为其不关注通过符号标记或通过声学特性的分析的音乐相似性。其也不同于标准的治疗方法,例如情绪的分类。

[0096] X- 系统通过 INRM(对音乐的先天神经生理反应) 的预测性、确定性模拟 (奥斯本 2009,未出版),如图 1 所示,以及通过接近身体和精神的目标状态的通路的构建来工作。B 部分详细解释了 INRM。简而言之,INRM 范例采取听觉的标准解释,从耳道到耳蜗的卵圆窗。模拟耳蜗本身以复制人类听觉特征的模型。范例进一步假定了到下丘和初级听觉皮层的神经通路。通过对镜像神经元 (mirror neuron) 和与前运动区相关的系统的简单模拟来预测脉冲和节奏性相关的唤醒的水平,包括节拍感应以及节奏功率和密度的指数。其它的音乐的生物活性特征也可以被模拟,其中包括例如在右前次级听觉皮层中的节奏模式的识别。

[0097] X- 系统附加地模拟黑索氏回 (Heschls gyrus)、后颞平面 (posterior planum temporale)、颞上沟和岛环状沟 (circular insular sulcus) 的活动,以预测与唤醒相关的音质和与指数级数相关的频率结构,包括八度音阶等值。还存在其他模拟的可能性,例如,利用例如谐度指数来模拟颞上回前部 (planum polare) 中色度 (chroma)(旋律的单个音符) 之间的与唤醒相关的效果。

[0098] 最后,计算“扰动”的一般水平以作为对核心情感位置和器官——例如导水管周围灰质和杏仁核——中唤醒和逆唤醒的预测。

[0099] 将计算出的预测性唤醒和逆唤醒的值结合,以模拟自主神经系统和例如 HPA(下丘 - 垂体 - 肾上腺) 轴这样的相关系统中的唤醒和逆唤醒的过程。

[0100] 传感器可选择地用于确定用户的唤醒状态,并通过 INRM 范例的预测性模拟分类的音乐被流动式接受 / 回放,以实现用户的目标唤醒状态。在一个可选择的实施例中未提供传感器。改为,初始和目标状态都是直接或间接地自选择的 (例如,通过选择具有关于用户真实的当前状态的唤醒值的“开始曲”)。例如,用户做出粗略的初始选择,首先,他 / 她最初可以从歌曲之间一一跳过,直到发现一个 (即通过反复试验) 既“喜欢”又“适合”他们的初始状态的歌曲。由此,在无传感器的实施例中,X- 系统可以根据预期的正常人的反应来创造趋向所需的唤醒状态的播放列表。

[0101] 另一替代方案中,提供了一个将一群人看作具有软件但没有传感器、依靠平均预期的反应的系统的实施例。一种应用是适用于“人群”的应用程序,其中,自动的电台音乐节目主持人 (DJ) 将能够在聚会中操纵一群人的情绪。

[0102] 其它的替代方案包括这些应用,即,通过传感器发送情绪暗示给系统来控制个人的音频环境,以及通过传感器或无传感器的输入来调查群体的情绪,以便诱导个人或群体接近所需的反应。

[0103] 其它可选的应用包括,根据音乐的神经生理内容的音乐的搜索、选择、说明、检测、分享或推广。

[0104] 在所有的系统和活动都涉及音乐和唤醒的情况下,个体之间的存在反应上的变差,以及由身体和精神、药物等的极端或罕见的状态造成的变化。X- 系统的优势是,其以最根本的生理反应为基础工作,其可以随着用户有意识和无意识的应答在道德和民主的协同作用中发挥作用。基于 INRM 分类系统的进一步的优势是,其可以适用于任何人类文化的音乐,并真正地适用于声音设计和自然世界的声音。

[0105] B. 对音乐的先天神经生理反应 (INRM) 的详细解释

[0106] 图 1 表示与听觉处理和解释有关的神经结构的简化模型。本发明的实例 X- 系统可模拟这些系统响应于以下部分所描述的声音 (例如音乐的) 刺激作出的运行和行为。

[0107] 对音乐范例的先天神经生理反应是精神和身体对音乐的最基本的反应的预测性、确定性的模型。虽然对音乐的反应深深地受文化、个人的历史和背景的影响，但是基本的神经生理反应通用到所有的音乐体验中。目前在神经生理学和神经科学中研究的真正的主体——包括来自于功能性磁共振成像、脑电图 (EEG) 和电子发射断层扫描 (Positron Emission Tomography) 的证据，以及与内分泌和自主活动相关的研究——使得建立大脑的下层皮质、皮质下和边缘部分如何对音乐起反应的预测模型成为可能。

[0108] X- 系统使用以下协议用于音频输入。输入取自于未压缩的 WAV 文件或任何其它合适的格式（当 X- 系统进行远程分类时——例如分类在远程服务器或个人装置上的音乐音轨，X- 系统可使用较低质量的文件格式。同样地，较高质量的文件格式可能更适合于其它情况）。如果音轨是立体声，我们通过将其均分来结合两个声道。这是特别重要的，例如，对于上世纪六十年代的音轨，其中一些响亮的乐器全部配置在左边或右边。这样应该不会导致干扰，除非音频由故障的立体声设备传出（例如错位的磁头）。音轨被分成给定长度的部分，并且独立地对每部分进行分析。

[0109] 图 7A 是表示在 X- 系统中用于分析谐度的主要组件的方框图，以及图 7B 是音乐分析工具的全部主要组件的方框图表示。主要组件的运行将在 B 部分的其余部分中描述。

[0110] B.1 耳蜗和初级听觉通路

[0111] 耳蜗和初级听觉通路的模拟通过使用 A- 加权滤波器来实现，如 IEC（国际电工委员会）61672 中所规定的。这弱化低频和增强高频，向着人类听觉的频率上限再次快速地下降；滤波器“拐点”在 6kHz 左右。需要用该加权来确保，（如在人类听觉中）高能量低频率的声音不会覆盖其它频谱信息。参照图 10。

[0112] B.2 谐度：黑索氏回和相关的音质图 (tonotopic map)

[0113] “谐度”描述了声音（例如音乐）与谐波列 (harmonic series)（当风吹过树洞，用你的手指轻轻地拨动小提琴或吉他的弦，或者渐进地吹长笛上更刺耳的单音符时，谐波列出现在你听到的声音中）的模式的对应。该谐波列在对称的共鸣对象中是声能量浓度的通用模式：基音 f，伴随其谐波 f2、f3、f4 等联合发出的声音。该模式在有感知的生命形式的进化过程中自始至终都是非常重要的，从原始细胞的谐波共振，经过感知在环境中的谐波声音的“安全性”，到乐器和人类声音的赏心悦目的谐波共振。“谐度”或者谐波列的模式的对应是通过黑索氏回来检测，其位于大脑的初级听觉皮层中。谐度激活大脑的核心情感中心中逆唤醒和快乐的中心。失谐度 (Inharmonicity) 或者缺乏与谐波列的对应激活唤醒系统。

[0114] X- 系统通过确定谐度和失谐度的水平来模拟黑索氏回对声音的活动和反应。这可能是一个复杂的过程。音乐结构可包括数个基音，每个都有各自的谐波或非谐波频谱。

[0115] X- 系统是前所未有的，因为其将音高和音色的所有情感处理结合成两个与谐度相关的算法。音色（声音的内部结构“色”），谐度（内部结构对应于谐波列的模式的程度）和个别的音高是在初级听觉皮层中进行最初地处理。处理音色的主要区域是黑索氏回和颞上沟，延伸到岛环状沟中 (McAdams 等 1995；Griffiths 等 1998；Menon 等 2002)。音高是逐渐地更加深入地在黑索氏回周围的区域中处理：色度（或者说八度音阶中音高的差异，例如在大多数传统的旋律中）激活黑索氏回和颞平面前部的双侧区域，而音高的改变（八度音阶移调等，例如男人和女人唱同样的曲调之间的差异）激活黑索氏回和颞平面后

部的双侧区域 (Brugge1985 ;Pantev 等 1988 ;Recanzone 等 1993 ;Zatorre 等 1994 ;Warren 等 2000 ;Patterson 等 2002 ;Formisano2003 ;Decety 和 Chaminade2003 ;Jeannerod2004 ;Talavage2004)。谐度和音高结构激活杏仁核和海马区域,并进而激活自主神经系统、核心情感中心、和内分泌和神经传递系统 (Wieser 和 Mazzola1986 ;Blood 和 Zatorre2001 ;Brown 等 2004 ;Baumgartner 等 2006 ;Koelsch 等 2006)。X- 系统通过分析垂直谐度窗口来预测性地模拟简单音色的神经生理的感觉 (黑索氏回、颞上沟、岛环状沟) :X- 系统通过谐波产物频谱 (harmonic product spectrum) 的计算来检测主要的基音,之后在不同的基音频谱内部和之间确定谐度的度。该分析“垂直地”应用于瞬时时刻,和“水平地”应用于音高和频谱随时间的演进 (与黑索氏回周围区域的音质映射相关) 并以线性谐波损失的方式表示出来。

[0116] 在一个非常简单的实施例中,线性谐波损失 (C) 的平均值和瞬时的谐度 (H) 结合在一起计算一段声音的失谐度 (I),其中 :

[0117] I = C/10-H

[0118] 该等式是如何计算失谐度非限制性实例,并且将 I 与 C 和 H 结合的其它方式可能也是适合的 ;而且,I 可以以其它或附加的变量来定义,就像 C 和 H。如图 12 和 13,表示随时间变化的谐波能量和损失。

[0119] 关于谐度计算的更多细节现说明如下 :

[0120] B.2.1 频谱分析

[0121] 首先,利用 8192 样本的窗口长度和 2250 样本的间距 (0.05 秒) 进行音频的 STFT(短时傅里叶变换)。其产生时间相对于频率的二维数组。

[0122] B.2.2 耳蜗建模

[0123] 在处理节奏的情况下,分析是在输入样本数据的变换的实例上进行的,其解释听觉通路的某些方面,主要是耳蜗拾音器 (cochlea pick-up)。耳蜗的特性是众所周知的,并且已经发展出了精确地模型。我们将基于频率的增益函数应用于输入信号,该输入信号减弱低音信号并放大高音部分,在在 6kHz 具有滤波器“拐点”。使用的确切变换是在 IEC61672 中规定的“A 加权”。

[0124] B.2.3 基本频率的检测

[0125] 在 STFT 数组的每个时间片中,基本频率的确定是利用谐波产物频谱的方法 (harmonic product spectrum method),如下所示 :

[0126] ●取频谱,通过因数 2、3、4 和 5,产生其沿着频率轴压缩的副本。

[0127] ●乘以所有 5 个副本 (包括原始的)

[0128] ●基本频率是由此产生的频谱的最大值。

[0129] B.2.4 平均谐度

[0130] 对于 STFT 数组的每个时间片,平均谐度是谐波能量和时间片中存在的当前总能量的比值。谐波能量是在后面的基音、以及基音的 1/2 和 1/4 的谐波中所获得的能量 :[1234567]。对于每个这些谐波,我们计算在最接近 STFT 存储桶 (STFT bucket) 加上每一侧的 3 个存储桶中获得的能量的总和,。

[0131] B.2.5 线性谐波损失

[0132] 在黑索氏回的周围区域 (颞平面、后颞平面) 的活动和演进 (progression) 的预

测——包括色度、八度音阶的变化和和弦进行（chord progression）等——被结合在一个单一操作中，描述为“线性谐度”或“谐波损失”。

[0133] 这完全是前所未有的：其根据每个步骤的偏离于谐波列的简单比例有多远来分析全部旋律和谐波的演进：线性谐波损失从STFT时间片上升，STFT时间片的基本频率与之前的时间片的基本频率不同。在基音上无变化时间片损失为零。基音频率第一次归一化是通过将其凑整至A440音调以下的最接近的音符值，之后，将其转换为单八度音阶。之后将该（归一化的）基音与前一个进行比较：如果它们完全一样，损失为零。如果新的基音是之前（归一化的）基音的以下谐波和次谐波（1/91/71/61/51/33679）之一，那么损失将定义为谐波的乘数或次谐波的除数。否则损失定义为15。

[0134] 线性谐波损失以每秒的损失表示。因此，度量标准既代表基音变化的比率，也代表变化的谐波距离。数值越大，表示更强的刺激效果。

[0135] 线性谐度将类似的情感系统激活为垂直的谐度（Wieser和Mazzola1986；Blood和Zatorre2001；Brown等2004；Baumgartner等2006；Koelsch等2006）。

[0136] B.2.6 谐度和效价

[0137] 垂直和线性的谐度都是效价的强有力的指数（Fritz2009），或者说声音是“正面的”还是“负面的”，“令人愉快的”还是“不那么令人愉快的”。线性谐度可追踪效价指标随时间的发展——原理只不过是谐波越多、正效价越多，谐波越少、负效价越多。

[0138] 可以想象的是，与黑索氏回相关的等式可以由不同的数学方法重新构建。颞平面函数能用任何不同的方式逼近是极不可能的。

[0139] B.3 节奏性：镜像神经元、听觉皮层和前运动区皮层

[0140] 人类对音乐节奏的反应涉及一组复杂的精神和身体系统的活动（Osborne1.2009；Osborne2009.2；Osborne2012.3），包括感知系统、蜗背侧核（dorsal cochlear nucleus）、下丘和脊柱系统（Meloni和Davis1998；Li等1998）初级和次级听觉皮层（Peretz和Kolinsky1993；Penhune等1999）、镜像神经元（Rizzolati等2001；Gallese2003；Molnar-Szakacs和Overy2006；Overy和Molnar-Szakacs2009）、运动前和运动皮层、基底神经节、前庭系统和小脑（Zatorre和Peretz2001；Peretz和Zatorre2003；Turner和Ioannides2009；）、自主神经系统（Updike和Charles1987；Iwanaga和Tsukamoto1997；Byers和Smyth1997；Cardigan等2001；Knight和Rickard2001；Aragon等2002；Mok和Wong2003；Lee等2003；Iwanaga等2005），以及最终的身体和核心情感的系统（Holstege等1996；Gerra等1998；Panksepp和Trevarthen2009）。其中一些可能特别涉及能够使编码在声音中的感知行为、生命力情感和活力再生的镜像神经元的激励，及其在在收听者的精神和身体中的表现形式。高能量的快节奏激活自主神经系统和内分泌系统中的唤醒，例如下丘脑-垂体-肾上腺轴。慢节奏激活逆唤醒。

[0141] X-系统根据每分钟的节拍数检测基本的、“默认的”节奏脉冲。经常在确定韵律的时候遇到困难，但是X-系统通过节奏事件的功率（power of a rhythmic event）随时间的累积来接近韵律结构的唤醒效果。节奏事件的功率定义为节拍前的能量与之后的能量的比率。在一个非常简单的实施例中，每分钟节拍的值（B）与节拍强度（S）的平均值结合以产生节奏性的值（R），其中：

[0142] R = √ B*S^2

[0143] 该等式是如何计算节奏的非限制性实例，并且将 R 与 B 和 S 结合的其它方式也可能适合；而且，R 可以以其它或附加的变量来定义。一般来说，R 可以是 B 和 S 的函数，但是最佳的关系将依赖于各种因素。如图 11，表示随时间变化的的节拍能量。

[0144] 关于节奏性的更多细节：

[0145] B.3.1 耳蜗的模型

[0146] 根据前面的解释，节奏的听觉感知是由传统的耳蜗模型预测的：随着音频输入，在输入样本数据的转换实例上执行所有的后续分析，其解释听觉通路的某些方面，主要是耳蜗拾音器。耳蜗的行为是众所周知的并且已经开发出了精确的模型。我们将基于频率的增益函数应用于输入信号，该输入信号减弱低音信号和放大高音部分，具有 6kHz 左右处的滤波器“拐点”。使用的确切变换是在 IEC61672 中规定的“A 加权”。

[0147] B.3.2 节奏感应

[0148] 通过节拍感应，利用具体校准的起始窗口预测性地模拟原始脊柱通路和运动前环(pre-motor loop)（包括基底神经节、前庭系统、小脑等）的激活，所有这些都与对节奏脉动的原始反应有关。

[0149] 当然，节奏性是模拟样本基本节拍以及其中的高阶韵律结构的参数。其首先通过利用频谱峰值通量的检测确定音符起点(note onset)来计算。这些起点之后用于产生和给大量的韵律结构假设打分。利用 Dixon 的方法[音频节拍追踪系统节拍根源的评价(Evaluation ofthe Audio Beat Tracking System BeatRoot)，新音乐研究杂志(Journal ofNew Music Research)，36(1)，39-50，2007]来产生、过滤和和评分备选的假设。除了所描述的方法，我们将该程序延伸至包括起点事件周围的频谱通量的量级，以便评价更高阶的结构。产生的假设利用同样的方法进行过滤和打分，终端输出包含样本的基本节拍的预测，根据预测的韵律结构加权节拍的次级输出，其中，低音节拍的重音节拍越明显，该值越高。针对所有得分高于给定的阈值的节拍假说，置信度值也表示为这些输出的分布方差。将该置信度值归一化以允许横向比较样本。

[0150] B.3.3 自相关

[0151] 节奏模式识别和保持（例如在颞叶的二级听觉皮层中）是通过自相似性/自相关算法(self-similarity/auto-correlation algorithm)（例如 Foote http://207.21.18.5/publications/FXPAL-PR-99-093.pdf)进行预测性地模拟。

[0152] 首先音频在重叠步骤中用汉明窗口处理；针对每个窗的功率谱的记录，每个窗通过 DFT（精细的傅里叶变换）的平均值计算。这些系数通过梅尔缩放(Mel-scaling)感知加权。最终，第二个 DFT 用于创造倒谱系数(cepstral coefficient)。高阶 MFCC（梅尔频率倒谱系数）被删除，留下了 12 个低阶 MFCC，在 100Hz 速度时形成 13 维的特征矢量(12 加能量)。这些数据之后经过矢量自相关，绘制在二维窗中，其中，x 和 y 轴随时间绘制音轨的展开。向上读数——例如从 x 轴上的音轨的第一瞬间上——的“亮度”区域显示相似点，和可能的韵律结构。

[0153] 点的分布密度也用在节奏诱导的唤醒的预测指数中（密度越大，唤醒越高）。

[0154] B.3.4 功率

[0155] 镜像神经元系统——除了别的之外，其检测“节奏的”活动的功率、轨迹和意向性——的激活通过节奏功率的指数预测性地模拟，该指数包括音量级、音量峰值密度、“低

谷”,或能量缺乏和运行能量的动态分布的计算。

[0156] B.3.5 音量包络分析(Volume envelope analysis)

[0157] 音量包络的计算为振幅数据的5毫秒的片的均方根。

[0158] B.3.6 音量级

[0159] 这是简单的在时间段内的平均的均方根级。

[0160] B.3.7 音量峰值密度

[0161] 每片(通常10秒)音量峰值的数量——如通过最小峰值间距=100毫秒的矩阵实验室(MATLAB)寻找峰值函数(findpeaks function)获得——乘以超过音量平均值的峰值平均高度,除以音量标准偏差。

[0162] B.3.8 音量微分峰密度(Volume differential peak density)

[0163] 类似于音量峰密度,但是采用音量的第一微分。

[0164] B.3.9 音量波谷长度

[0165] 音量比音量平均值低标准偏差的一半的平均持续时间。

[0166] B.3.10 音量波谷的最小值

[0167] 音量波谷的音量最小值的平均值除以音量标准偏差。

[0168] B.3.11 动态分布图

[0169] 此外,在起始之前和之间的能量损失的分布图(陡峭对于高唤醒,平滑对于低唤醒),其出现重要的镜像神经元信息,将在以后通过导致重要的发音的能量流的分布计算来预测。

[0170] 例如,τ“tau”(希腊文第十九个字母)耦合(Lee2005):τx=K x,gτg

[0171] 其中,tau=滑音原点的时间(前一个起始的终点),x=在下一个可预测的起始之前的间隙,g=通过神经元组装的电能的模式流(patterned flow),kappa=由大脑确定的移动值。能量的分布图将由kappaXG的平均值的分布图确定。

[0172] B.3.12 标准的、市售的用于节奏检测的软件可能令人满意地用于一些音乐流派,但是这样的软件可能不能检测任何给定的音乐片段的特定生物激活的节奏,并且可能在检测某些节奏上依然有些困难。预测性地模拟大脑的核心节奏处理中心的激活的上述算法已经被证明是可靠的。这些算法中的一些——例如节拍检测——在理论上能被其它的数学程序代替。本发明的独创性在于前所未有的仿生性质。因此,我们在音乐(节奏)中有一个现象,就是已知对自主神经系统(也有核心情感系统、内分泌活动和神经传递)中的唤醒和逆唤醒产生效果,已知其进而对你的感觉产生强大的影响:放松、能够专注、想跳舞等。我们也具有测量节奏效果的的手段(我们的传感器)。我们的分类算法(上述的)作为来自于数字签名分析的有关数据的输入,并且产生对挑选的生物计量有预测性影响的输出。强烈的节奏将具有唤醒的效果而温和的节奏将具有镇静的效果,并且基于同样原理,没有现有技术中的不足。在模拟对节奏的先天神经生理反应中,假设、测试和细化将节奏的测量与其预期的对心率和皮肤电导产生的效果(在本实施例中)连接的算法。

[0173] B.4 扰动和核心情感系统(位置和器官)

[0174] 一段音乐的“扰动”涉及其在一段时间内变化的速度和程度,根据节奏性和谐度以及声压中的一般波动。

[0175] “扰动”将涉及上述通路的节奏和谐度的变化指数与听觉脑干和皮质活动结合,

该听觉脑干和皮质活动使杏仁核、海马和核心情感区域受神经支配，该杏仁核、海马和核心情感区域影响神经传递和内分泌系统，包括肾上腺轴、多巴胺回路和，例如去甲肾上腺素、褪黑素和催产素的水平 (Miluk-Kolasa 等 1995 ;Gerra 等 1998 ;Kumar 等 1999 ;Evers 和 Suhr2000 ;Schneider 等 2001 ;Blood 和 Zatorre2001 ;Grape 等 2003 ;Uedo 等 004 ;Stefano 等 2004 ;Herbert 等 2005 ;Nilsson 等 2005)。唤醒和逆唤醒的这一重要的预测器可被表示为节奏和谐度的微分。

[0176] 因此，"扰动"是音乐体验中的变化率的测量和变化程度的测量。这些因素似乎可以激活大脑的核心情感系统，例如杏仁核和导水管周围灰质，其进而与自主和内分泌系统连接。高水平的音乐能量扰动可提高唤醒 ;在低水平的音乐能量扰动可增加逆唤醒效果。

[0177] 一段的总扰动 (T) 被确定为片段的谐度 (H') 和音轨 (P) 的音量峰值期间存在的能量的扰动的结合。谐度的扰动计算公式为谐度的微分的标准偏差除以微分的平均值。

[0178] 在一个非常简单的实施例中，总扰动的计算公式是 :

[0179] $T = dH/dt*P$

[0180] 该等式是如何计算扰动的非限制性实例，并且将 T 与 H 和 P 结合的其它方式可能也是适合的 ;而且，T 可以以其它或附加的变量来定义。

[0181] 如图 14 和 15，表示随时间变化的音量和谐波能量。

[0182] B.5 结合值

[0183] 上述通过测试假设和完善的每个算法有效地成为了大脑的"虚拟器官"，这有助于我们预测对唤醒和逆唤醒模式的水平所产生的效果，其能够在音乐中利用数字签名分析来检测。每个"器官"的相对加权可以适于利用启发式的、机器学习或其它技术来校准一组协调一致地工作的"虚拟器官"的整体预测能力。

[0184] 上述分析的任何子集可以结合在一起以产生单数估算，其中一段音乐 (或其部分) 位于从放松至兴奋的尺度上。执行该结合的公式可能来源于实验数据，如下 :许多收听者听相同选择的音轨。每个收听者之后独立地将所有音轨排列成序，从他们认为最令人放松的到最令人兴奋的。(该排序过程也可以通过测量收听者的生理数据来客观地实施，但是，这到目前为止给了收听者很多不太一致的结果。) 统计回归分析在此之后执行，平均的人类排序作为因变量，以及音乐分析的所选子集作为自变量。换句话说，产生使用分析来预测人类排序的单一公式。在该公式中的系数被挑选以用于给出最好的预测，并针对所有音轨。所得到的公式可之后用于产生针对大量音轨的大规模地自动预测。考虑下面的示例数据 :

[0185]

音轨	平均的人类排序 (0-1)	谐度平均值 (mh)	音量级别 (vol)	节奏性 (rhy)
1	0.2	0.212	0.010	118
2	0.4	0.231	0.069	228
3	0.5	0.204	0.123	187
4	0.6	0.225	0.294	130
5	0.8	0.173	0.163	155

[0186] 任何统计回归方法可用于产生整体公式。例如,如果我们使用具有普通最小平方法估计的多元线性回归,我们将获得以下公式：

[0187] 预测的顺序＝ -6.59*mh+1.63*vol+0.0018*rhy+1.36

[0188] 也可以使用一个变量的非线性变换(例如对数或倒数)或者多个变量的非线性组合(它们的乘积或比例),这是通过预先计算它们以及之后将它们作为回归分析中的附加变量来实现的。

[0189] 用在每个算法中的系数和算法的相对加权相结合,可利用元数据(例如流派和艺术家)针对不同的音乐风格进行优化,该元数据通常携带在以数字化格式散布的音乐中,例如光盘和因特网上。随着大量的(匿名的)人类反应数据的累积,其可以反馈(经过收听者的同意)到网络部署的X-系统中,这将有可能微调两个方程的系数的相对加权并将它们的相对权重结合来提高精确度。类似的系数和加权的优化将通过将分析用户数据与音乐元数据(例如流派和艺术家)结合来实现,该元数据典型地在以数字格式散布的音乐中可用,并在适当的时候,这种优化将扩展到个别用户和特定唱片中。

[0190] 为每段音乐计算的全部唤醒指数可以表示为描述从开始听到结束的所有神经生理效果的单个数字,或者其可以以唤醒指数为纵轴和以时间为横轴地以图形表示出来。最终的轨迹将有效地描述收听者从开始听到结束的预期的收听者神经生理学行程。后者很可能特别用在更长更复杂的音乐片段中,例如大部分的经典曲目,反之一些其它的曲目,例如现代西方流行音乐可能更方便用单一数值表示。不论哪种情况,音乐片段的效果是固有的(因为,它是在音乐中检测到的模式的结果)并且依赖于收听者的状态(因为,音乐的神经生理效果是相对的而不是绝对的[Altshuler“I 情绪同态原理(The Iso-Moodic Principle)”1948])。

[0191] 当我们更深入更细致的学习诱导大脑时,以及当传感器技术进一步发展时,将开发不同的等式来预测不同的音乐结构在不同的可测量的输出上产生的效果。所有这些对音乐的先天神经生理反应的应用实例将作为本发明的不同的实施例,本发明要求保护一种预测来自于任何音乐传统的任何音乐片段对普遍的人类神经生理产生的效果的新的系统和方法,这通过分析音乐中的生物激活模式和利用适合于特定生物计量指数的数学公式来预测这些音乐模式在挑选出的生物计量指数上产生的效果来实现。

[0192] B.6 本部分描述一种替代方法，其是将节奏性、失谐度和扰动的值结合而获得兴奋值 (E)。在该变换方法中，E 由以下公式给出：

[0193] E ＝ (10*I*R)+T

[0194] 该等式是如何计算兴奋值 E 的非限制性实例，并且将 E 与 I、R 和 T 连接的其它方式也可能是适合的；而且，E 可以以其它或附加的变量来定义。

[0195] 这通常产生一个来自于 -1 至 7 之间的数字，表示逆唤醒 - 唤醒的数值范围。一般，五种类别的唤醒的阈值近似于：

[0196] -1 至 0.6 ＝ 1

[0197] 0.6 至 2.2 ＝ 2

[0198] 2.2 至 3.8 ＝ 3

[0199] 3.8 至 5.4 ＝ 4

[0200] 5.4 至 7 ＝ 5

[0201] 另一种选择是，节奏性和谐度相乘再加上扰动的等式。在一个实施例中，对数标度和斐波那契数列可用在听觉数据的分析中。

[0202] 详细信息：针对每个 R、H 和 T，X- 系统记录进一步分类成上升、下降或恒定的单个平均值（μR、μH、μT）和变化属性（ΔR>0，ΔR<0，ΔR ＝ 0；ΔH>0，ΔH<0，ΔH ＝ 0；ΔT>0，ΔT<0，ΔT ＝ 0）。

[0203] R、H 和 T 的平均值映射到（在最简单的情况下取归一化的平均值）表征生理状态的 n 维的点 p 上。R、H 和 T 的变差也映射到（再次，在最简单的情况下取归一化的平均值）另一个表征这些值在生理状态上产生的定向效果的 n 维的点 q 上。

[0204] p 和 q 的连结允许每个音乐片段被映射到音乐效果矩阵 M、一种 2*n 维矩阵上，n 维对应于由 E 所测量的、表示 E 可落入的粒状范围的生理参数，其它的 n 维对应于音轨对生理参数产生的效果（上升、下降或保持任何给定的生理参数或 E 的维度）。

[0205] 我们现在更详细地描述音乐效果矩阵 M 是如何产生的。如前面所指出的，图 7A 是表示 X- 系统中用于分析谐度的主要组件的方框图，以及图 7B 是音乐分析工具的所有主要组件的方框图表示。由分析输出的值指定为 t 中的函数，特定测量的时间指数。这些值（对应于 R、H 和 T) 分组如下：

[0206] X(t)：关于节奏“存在”的值，与脉搏相关的节奏结构节拍、功率和密度，和与大脑皮层活动、核心情感位置，和自主和内分泌反应有关的谐波节奏的值。

[0207] Y(t)：在人类感知的极限内，与旋律与谐度中的与指数系列相关的频率结构符合度——涉及耳蜗、黑索氏回和大脑皮层处理，核心情感位置和自主和内分泌反应。

[0208] Z(t)：X(t)、Y(t) 和动态功率 (W(t)) 中变化的比例和幅度，其测量是利用归一化的、增益调整的音量级别——涉及核心情感系统、和内分泌和自主神经系统的变化。

[0209] 在分类之前可以是关于音轨的聚集、记录出处、流派和其它数据。这可能是根据行业标准来执行的，例如如Gracenote®提供的行业标准，这可能是个人用户编辑的、众包方法 (crowd-sourcing method) 的结果，例如协同筛选，或者可能是未来聚集标准的结果，其基于，例如数字签名分析。聚集的目的是允许用户挑选喜欢的音乐风格，虽然这对于 X- 系统的正常工作不是严格必须的。

[0210] 为了减少分析一段音乐的计算费用，只检查某些区域。基于递归地检查具有大的

变化率的区域的配置参数和适应机制，这些区域的位置和长度被动态地确定。这产生了针对每个函数的稀疏数组值，由时间指数确定。由于递归分析，步长 t 在函数定义域 t 中将有所不同。

[0211] 在算法上，这些区域通过将窗口函数 (windowing function) 应用于引入的音频数据来产生。之后，采样窗口在该区域“分级”，并且将每级的结果聚集以形成在时间 t 的单个输出。例如，一个区域可包括 (绝对的) 时间间隔 (0s ;1s)，其进一步加窗为 50ms 的样本，其具有 10ms 步长。这产生了总共 96 个样本点，将其结合以形成单一值 X(0) = x。

[0212] 由“声流”执行 X(t) 的分析——基于将节奏感应与模式识别和功率和密度指数相结合。

[0213] 利用两项主要的技术执行节奏感应 ;带限的功率谱密度起始分析、自适应梳状滤波 (adaptive comb filtering)。之后，两个技术的结果经受一定数目的基于音乐理论的探试法，并结合以形成音乐节奏的单一估计。

[0214] 探试法包括例如将合理的节奏最小化和最大化的规则，或者适合于如果已知的给定输入流派的节拍的一些概率分布。其也可以包括基于输入的某些频段的加重和去加重。

[0215] 频谱密度起始分析使用加窗样本的短时傅里叶变换序列来计算存在于特定频带中的能量。该数据被暂时追踪以观察频带中的峰值，其表征节奏事件。

[0216] 梳状滤波包括输入信号的卷积，其具有各种不同间距的脉冲序列，在脉冲间隔接近输入节奏的基础上，整体卷积结果将增加。该技术在之后递归地用于发现表征输入节奏的最适合的脉冲间隔。

[0217] Y(t) 值的确定是通过听觉场景分析的改编。音频输入数据经过伽马音耳蜗滤波器组 (gammatone cochlear filter bank)，将其分成多个流。针对每个流，计算特设的频率和起始信息。

[0218] 从每个流的立体声音轨获得空间信息，频率峰值利用傅里叶变换来计算，起始探测器映射 (onset detector map) 用于寻找声音元素的开始。

[0219] 该信息被合并并且被相互关联，以分割输入到声源中的音频数据。针对这些声源中的每一个，数字的计算公式为其基频的谐波内的声能与其基频的谐波外的声能的比值。Y(t) 是针对每个片段的声源的比例的平均值。

[0220] 基频的确定是利用谐波积谱 (Harmonic Product Spectrum) 实现的，其中，信号反复乘以其自身向下采样的副本，导致在频谱中出现对应基频的大的峰。标准信号处理技术也适用于将产生的输出去噪。

[0221] 在 X(t)、Y(t) 和 (W(t)) 中的变化的比例和幅度下测量 Z(t)。

[0222] 在每种情况下 (X(t)、Y(t) 和 Z(t))，系统记录单个平均值 (μX，μY，μZ) 和进一步分成上升、下降或稳定的变化的曲线。

[0223] ●上升的——在函数 X(t)、Y(t) 和 Z(t) 中是整体正趋势的。

[0224] ●下降的——在函数 X(t)、Y(t) 和 Z(t) 中是整体负趋势的。

[0225] ●平稳的——仅微小地偏离于音频输入信号中产生的平均值 μ。

[0226] X、Y 和 Z 的平均值被映射到 (在最简单的情况下采用归一化的平均值) 表征生理状态的 n 维点 p 上。X、Y 和 Z 的变差也映射到 (再次，在最简单的情况下采用归一化的平均值) 另一个表征这些值在生理状态上产生的定向效果的 n 维的点 q 上。

[0227] p 和 q 的连结允许每个音乐片段被映射到音乐效果矩阵 M、一种 2*n 维矩阵上，n 维对应于由 E 所测量的、表征 E 可以落入的粒状范围的生理参数，其它的 n 维对应于音轨对生理参数产生的效果（上升、下降或保持任何给定的生理参数或 E 的维度）。

[0228] C. 如何使用 X- 系统

[0229] 如上所述，X- 系统可使用受检者的生物计量数据（传感器是可用的情况下）来测量神经生理唤醒。之后，其引导受检者分阶段接近这种目标的唤醒水平、精神和 / 或情感状态。这是在具有音乐数据库的情况下实现的，其利用先天神经生理反应的预测模型进行预先分类。也可能是实时或近实时的分类。分类可以直观地显示出来（例如在用于音乐回放的计算装置的显示屏上）；这包括针对每个音轨的 E 值的显示，或 E（兴奋值）值在音轨中如何变化；R、I、H、C 和 T 可以直观地显示出来。选择预测或匹配受检者当前的神经生理唤醒水平的一段音乐，并且在音乐每部分的基本音乐效果的基础上创建播放列表。听播放列表会引导或鼓动用户通过随着音乐的无意识的神经生理诱导接近期望的唤醒水平、精神和 / 或情感状态，并使这种水平能够保持下去。受检者当前的神经生理唤醒水平也可以直观地表示出来，能够收敛到期望的目标状态。

[0230] 在一个实施例中，X- 系统是设计用于感知用户的精神和身体状态，和成流所选曲目以实现目标状态，例如：

[0231] ●兴奋

[0232] ●放松

[0233] ●集中

[0234] ●警觉

[0235] ●身体活动的强化

[0236] 例如，参照图 2、3 和 9。

[0237] C.1X- 系统中的组件

[0238] X- 系统包括：

[0239] ●能够根据特定的唤醒和逆唤醒水平将所有文化的音乐远程或近程分类的自动分类软件；这种分类可被提供用于一般独立使用的传感器和诊断软件。这可能是基于奈杰尔·奥斯本的 INRM（对音乐的先天神经生理反应）范例。

[0240] ●手动或自动分类的音乐数据库（利用自动分类软件），以实现特定的唤醒和逆唤醒水平

[0241] ●传感器，检测唤醒（例如激动）和逆唤醒的生理指标（困倦），包括心率和皮肤电导

[0242] ●诊断软件，其使用传感器数据，以监测用户中的唤醒和逆唤醒水平

[0243] ●音乐回放 / 成成流（例如播放列表选择）软件，其选择从数据库中预先分类的音乐来成流适当的曲目，以通过逐步的诱导过程实现精神和身体的目标状态，该过程开始于当前诊断出的“状态”；通过诊断软件监测接近这些目标的进度。根据来自于收听者的生物反馈，可选择针对收听者的特定音轨用于回放（通过成流或其它方式）；播放列表可以在本地创建并且要求的音轨用于成流 / 下载等；这也可能用于将生物反馈和期望的“状态”信息发送给远程音乐服务器，并且为此，服务器产生适当的播放列表且提供音轨给本地、个人的回放装置。在该变体中，个人回放装置不必有本地音乐库或 X- 系统软件 / 固件等；其仅

需要有检测收听者的音频偏好和生物反馈数据，以及利用低容量的反向通道将其传递给远程服务器，并且之后接收来自于远程音乐服务器的音乐的能力。

[0244] 需要注意的是，所有软件也可以实施在硬件、固件、系统级芯片中，作为第三方音频堆栈的一部分，以及以其它任何方便的方式。

[0245] 附录 1 是 X- 系统组件的更详细的说明。

[0246] C.2X- 系统的实际应用

[0247] 传感器意在用于测量一个或多个预先确定的用户的精神和身体状态的参数，并将该信息传递给处理器；该处理器设计用于从音乐分类数据中选择适合于将用户从她的 / 他的当前的精神和身体状态引导到唤醒和逆唤醒的预期状态的音轨。该组合将允许 X- 系统：

[0248] ●实时地感知人的精神和身体的神经生理状态；

[0249] ●分析用户的音乐收藏，或他 / 她有权访问的任何其它的收藏，例如以音乐服务为基础的基于云或远程 / 中央服务器；和

[0250] ●根据期望的唤醒状态计算和传递播放列表。

[0251] 这将使用户诱导自身到期望的状态，例如：

[0252] ●被激励并准备运动或锻炼；例如，为了比赛而增加血氧水平或减少手术后的恢复时间；

[0253] ●使自身放松并能够渐渐入睡；

[0254] ●在深思状态中支持洞察力的发展；

[0255] ●在深思状态中支持创造性思维的发展；和

[0256] ●保持专注并能够集中精力。

[0257] （例如，提供支持来克服例如失眠的状态，以减少创伤后精神紧张性精神障碍(PTSD) 和狂躁症患者中的药物治疗，从而开发和组织记忆，通过数据保留所需的短期、中期和长期来分类。），并创造一种鼓励创造力和想象力的状态。

[0258] 图 4 的图表示 X- 系统的当前计划的实施例。在一个可选择的图 4 的实施例中，因为无处不在的移动计算模糊了装置之间的差别，图 4 中示出的元件在用户的 PC 内（音乐播放器、音乐库、自动化分析和音乐效果数据库）可以分布在两个或两个以上的计算装置中。在一个商业实例中，其也可以设置成与便携式音频装置一起工作：参照图 5。

[0259] 虽然这些组件是 X- 系统的关键元件，其核心创新技术是音乐的生物活性成分的定义（根据预测性的对音乐的先天神经生理反应的范例，Osborne2009，例如参照图 1），该算法用于根据数字签名分析和校准方法来计算这些生物活性成分，该校准方法用于使系统和个体的先天神经生理反应一致。

[0260] D. 一个或多个传感器

[0261] 传感器可以是腕带、手持或任何其它适于执行必须的参数测试的形式。传感器可以是身体搭载的，或使用耳机的（例如，将传感器合成到头戴式耳机中），通过红外或听觉、无线，或更普遍的任何形式的生命传感来远程监测的。捕获到的数据优选包含生物计量参数，例如，心率（包括脉搏节律分析）、血压、肾上腺素和催产素水平、肌张力、脑电波和皮肤电导性。可选择的设备形式包括项链、手镯、嵌入在衣服中的传感器、其它首饰、皮肤下植入的传感器、头戴式耳机、听筒、手持式的传感器，例如涵盖了手机、MP3 播放器，或其它的移动计算装置。

[0262] 当前用在X-系统项目中的传感器包含用于测量皮电反应(GSR)腕带传感器,和用于心率和血氧测量的标准夹指氏血氧仪。为了商业化的目的,这些传感器将结合成单一的、可佩戴的无线装置。其它的潜在的生物传感器和移动传感器也可以包括在内,只要它们在经济上可行。

[0263] 这些传感器必须能够与任何其它可能的测量结合来测量脉搏率和皮肤电的组合体,以及必须对用户的运动或环境的变化的破坏有抵抗力;其同时必须使长时间穿戴该传感器没有不适和尴尬成为可能。其它的传感器包括实体的生物传感器,例如氧合作用、EDA、EDC、EDR、ECG、血糖水平、BPM、EEG等,以及多频谱传感器(无线电、红外、紫外、热力学、和广谱),其检测身体辐射光环。

[0264] 图5表示期望的体系结构概述。图5表示本发明的X-系统的一个实施例,其中,基本的音乐库和分析软件存在于用户或第三方远程或本地地可操作个人电脑中,具有将选择的音乐传递给个人音乐播放器装置的能力,之后,其基于可用的音乐产生动态的播放器列表。

[0265] X-系统传感器测量某些挑选出的用户生理状态的参数并将结果数据无线传递给(或与之通信)播放列表计算器的处理器,其存在于或是以其它方式连接到音乐回放装置的(例如,个人电脑、智能手机、MP3播放器或其它音频装置)。传递方式优选无线地,但是需要理解的是,其它传递方式也是可能的。当然,处理器可以与传感器集成。

[0266] 挑选出的生理状态参数由P表示。函数F(P)将这些参数减少为单个的、归一化的点E,表征用户的一般生理状态。在最简单的情况下,E是用户的生理唤醒(或逆唤醒)的一维测量。随着进一步的输入,可能会获得更加复杂的测量,从产生n维的点E。一种有效的原型已经被开发出来,其利用脉搏率'p'和皮肤电导性'v'来计算生理唤醒的简单指数:其中$E = p+v$。目前,该原型使用的是美国燕牌指夹氏脉冲血氧仪(Nonin X Pod Pulse Oximeter)和皮肤电导生物传感器。用户的脉搏率、血氧和皮肤电导是不断地监测的;心率可用于控制电传导中的平均值。这两个传感器目前都是独立地工作并且与控制计算机无线连接。其可以被替换为一个集成的传感器。可选择地,采用任何其它形式的有线或无线连接,将传感器输出与播放器与输出装置连接都是可能的。附录1给出了更多的细节。

[0267] 用户最初提供具有其个人音乐收藏的系统(或者使用可成流或可下载的线上音乐库)。这是针对兴奋水平的分析,其利用INRM分类与信号处理和机器学习技术相结合。用户之后使该信息与其音乐播放器同步并选择兴奋/唤醒的水平;除了用户的其他人也可以选择兴奋水平。腕带式传感器提供给系统不断更新的用户兴奋度的实时状态,允许该系统对在用户中产生的外部影响做出反应,并且"捕获"它们,利用诱导的原理将它们带回到接近期望的状态。一旦用户已经达到兴奋的目标水平,他们将通过确定有效于保持该状态的音乐被保持于此。

[0268] 虽然X-系统的传感器的当前版本是基于心率和皮肤电导,但是存在针对其它测量的前期整合的强有力的论据,该其他测量包括例如EEG、脑电波传感器。这将允许例如集中、警觉、沉思、困倦或创意流这样的因素直接被监测,这是通过大脑中的神经元诱发启动的频率的感测,而不是直接通过唤醒的指示器。第二组相关的激发存在于机器学习的进一步的方面中。个体的生理反应差别很大,因人而异,取决于当时的情况、新陈代谢状态等。X-系统可从个人用户得知他们的生理反应的范围,以便识别相关的唤醒水平,并单独校准

诊断软件。正如已经阐明的,其也可以通过他们所选择的曲目得知他们的个人喜好。X- 系统也可以直接从一组音乐特性出发,利用神经网络来预测这些特性在生理测试上产生的效果,不是先减少这些特性到预期的兴奋 / 唤醒水平。

[0269] E. 音乐选择算法

[0270] 某些神经生理唤醒的水平一些活动的必要的先兆,例如,睡眠、放松、加速学习和研究,或提高警觉性和活动力。用户将优选地被呈现以用户界面并从这些活动菜单中挑选,以便使系统建立唤醒和情感的目标水平,这将有利于所选的活动。

[0271] 表征由传感器诊断出的受检者的神经生理状态的点 E 用于从音轨数据库中选择音乐,该音轨数据库是由音乐效果矩阵 M 索引的,基于粒状点 r 和指向生理状态的方向 d,用户已经选择移动向该生理状态(更多细节可参照之前的 E 部分)。

[0272] 选择的第一段音乐将对应受检者的初级神经生理状态,由 E 表示。随后的片段是基于它们在 M 中的值选择,这样,每个都将——按顺序播放——有能力逐步地引导受检者的状态接近目标状态。一段音乐被选择包括在播放列表中的顺序由矢量确定,其代表 M 中合适的、系列的音乐效果值临时组织的上升,或下降。数据库中满足该系列效果值的要求的一组音乐片段被称为‘合格的目录’。

[0273] 该合格的目录根据一组规则安排在实际的播放列表中,包括但不限于随机选择、防重复、流派偏好或一些其它的探试法。在某些情况下,可能要遵守 US 数字千年版权法案(DMCA)。

[0274] 在使用传感器的情况下,之后会建立生物反馈回路,以便保证播放列表连续的重算来校正基于全部情感的任何维度的干扰、个体敏感性和其它因素,其易受连续测量的影响。接近非测量的精神和 / 或情感状态参数的方向将依然会发生,尽管缺少生物反馈回路,因为神经生理唤醒是精神状态和情感的必要的前提,并且其建立了一些条件,在这些条件下,收听者最容易受到所有音乐效果的这些其它方面的影响。

[0275] 一旦已经播放一段音乐,为了最小化循环周期,优选将其从潜在可用的目录中移除,以便避免不必要的重复播放。该防重复规则服从可行性试验,以便可以将合适的严重性消息显示给用户,如音乐数据库中内容或内容多样性不充足的警报,以使系统能够随着改进的建议有效运行,例如更多音乐片段的推荐,可以将其添加到数据库中以改进其运行。

[0276] 在该情况下,目录已经被分散式的预分类或者从中央服务器中成流,播放列表可以以动态模式进行初始化计算,该模式中,从数据库取用短片段。一旦收听者已达到唤醒的目标水平,长片段将被允许进入到合格的目录库中,为的是播放列表计算和系统可进入到维持模式。由超过一个预先确定的因素引起用户的唤醒水平变化的任何干扰,可导致系统重新进入基于短片段的动态模式和重新计算播放列表,以便诱导用户以更快的速度回到目标状态。

[0277] 适用于短片段的防重复规则可用于计算最小要求的目录尺寸,其以独立的音乐风格的数量为基础,该独立的音乐风格数量可以由用户、短片段的平均长度、最小周期数、以及落入音乐效果矩阵的最小构成单元中的可用的短片段的数量来选择,在防重复规则允许歌曲或片段回到选择池中之前必须通过该最小周期数。

[0278] F. 音乐播放器

[0279] 音乐播放器可以是标准行业软件的改编,例如 Windows 媒体播放器,其能够根据

音乐选择算法建立动态的播放列表，以及能够提供给用户更多的辅助程序，例如，音乐风格的选择、相关元数据的显示和视频目录。

[0280] 音乐播放器也可以是从通过互联网访问的应用软件商店下载的应用软件。图8概括了一种播放器系统和与传感器子系统集成的设计。在一个实施例中，播放器系统和子系统可以分布在两个或两个以上的计算装置中；与移动计算和个人输入相似的普遍存在的计算方法可以与处理和传递音频输出的多种方式一起使用，无论是私人的还是公共的。因此不仅播放器，而且处理器和人机互交装置，包括但不限于通过情感暗示以及排序或定序消耗的交互作用的诱导和个人环境的控制，可以用在实施方式中。

[0281] G. 诊断和成流软件

[0282] 当传感器用在X-系统中时，那么诊断和成流软件就有读取来自于传感器的值和确定用户的唤醒状态的能力。皮肤电导的性质意味着绝对值能够明显地改变，这是因为其与皮肤的接触程度，因人而异并通过正常出汗实现的。为了解决这个问题，皮肤电导值可以基于用户的心率进行自动校准。

[0283] 系统的用户佩戴该系统，选择他们喜欢听的音乐曲目，决定他们想要达到什么兴奋水平并穿上传感器。一旦对用户的唤醒状态作出诊断，随着选择的兴奋水平，该数据会用于选择来自于曲目的音轨程序。

[0284] 可选择地，用户选择音乐的曲目，例如，爵士乐、古典的、印度的、世界、巴洛克式，并决定他们应该达到什么唤醒水平（例如放松、兴奋、稳定）并穿戴该传感器。一旦对用户的当前唤醒水平作出诊断，曲目将自动选择引导或“诱导”收听者从其当前状态进入到他们所选择的唤醒状态。这是通过定义播放列表执行的，其诱导用户从由INRM参数定义的多维空间中的当前情感位置，以小步长移动至接近INRM空间中已经定义的位置，其由期望的终点定义。

[0285] H. 手动分类

[0286] 在一个实施例中，曲目已经被手动地分类，这是利用节拍器通过脉搏/韵律的综合检测，以及关于唤醒和逆唤醒的水平的直观地预测判断，该唤醒和逆唤醒水平与各种音乐参数关联，包括节奏、谐度、扰动等，例如，脉搏/韵律越快，唤醒越高，谐度越高，唤醒越低。在图16的样本分类中（选自迈尔斯·戴维斯的曲目），音轨被置入到对应激活/唤醒水平的五个类别之一。

[0287] I. 手动分类矢量

[0288] 举例说明，在另一种手动分类中，音轨被进一步分成平稳、上升和下降的矢量，例如，如果用户选择高的激活/唤醒目标状态，“类型4上升”将被选择；如果用户希望保持适度激活状态，“类型4平稳”将被选择。例如，参照图17。

[0289] 在图18的实施例中，选自贝多芬交响曲的乐章已经根据矢量进行了分类。需要注意的是，没有乐章被识别适合于4/平稳或2/平稳。

[0290] 本发明的实施例已经参照其对人类的影响而被说明。然而，音乐对动物的影响是有据可查的。这几乎确定取决于声音环境的简单心理声学影响，而不是就音乐/生物本身的论述，但是除了一般的消费者、专业的、运动的、健康、医疗、以及其它市场，本发明的实施例还可看到在畜牧业或兽医学中的应用。

[0291] J. 社交网络

[0292] 在本方面的应用中，X- 系统作为诊断、监测或诱导工具，适用于促进由 X- 系统的算法确定的神经生理状态、唤醒、情感和效价数据通过短距离无线和蓝牙网络和更广泛的社交网络——例如，Facebook 和 Twitter——传递给朋友，以及传递给医护人员。

[0293] 该应用实现了一系列在智能手机和其它装置上的引导和连通的应用程序，能够允许用户‘连通和引导唤醒的先天状态’（情绪或情感）以及‘连通和引导体验’。其使个别的 X- 系统用户不仅能够看到显示器显示他们自己的先天状态，而且，当他们体验丰富多彩的活动时，从听音乐到运动和在医疗机构中的恢复和术后护理，允许其他人‘读取’他们的真实的或无意识的状态。

[0294] 一种系统和方法，其用于连通 X- 系统诊断能力，以解码神经生理状态，将其改编，以促进更深入的、更直接的唤醒状态和效价的连通，同时将其用于社交网络中的个体和群体之间的各种各样的活动（包括但不限于音乐）。

[0295] 一种根据唤醒的实际状态（由 X- 系统测试的）产生请求信息给搜索引擎的系统和方法，该搜索引擎例如 Google——该唤醒信息能够在之后作为给搜索算法和控制显示的广告的算法的输入（例如，当在特定的唤醒状态时网络用户可以更容易接受关于某种产品的广告，并且利用唤醒状态信息将搜索结果和广告定制为最大相关性）。该唤醒信息也能够用于指示‘当前’状态信息（“我心情很好，正在听贝多芬”等。）。

[0296] X- 系统将个体的先天神经生理‘状态’按照数据的连续体和精细的类别进行分类，从 1（高唤醒）到 5（逆唤醒）进行排列。这将核心的 X- 系统应用连接到了音乐选择上。

[0297] 在该‘社交网络’或‘分享’应用中，先天的‘状态’唤醒 / 逆唤醒和个体的效价数据被各种标准的通信网络（包括但不限于 Wi-Fi、蓝牙、GSM，以及其它移动网络和固话互联网）直接地以及经过广泛社交网络系统（例如 Facebook）传递出去，以使端对端和一个对多个的唤醒连同（可选地）指示同时的音乐或其它的娱乐选择，或自发的活动（‘正在看电影的是我；响应一个广告；在村庄中行走；跑步、骑车’）的编码的连通成为可能，所有都是实时的，或接近实时的。例如，X- 系统检测音轨的情感唤醒参数信息，并且之后将该信息嵌入到音轨中或与音轨连接的电子设备中，或者作为与音轨关联的元数据。

[0298] X- 系统‘状态’数据能够被分成实时快照（当前的唤醒和效价）；实时流（连续流）；作为历史（昨天的唤醒和效价），随着或不随着当时所选择的音乐的数据。这可能被称为“个人的活力指数”（活力：有生气；活泼）。

[0299] 该数据之后将作为图形、彩色代码，或各种统计表方式被显示出来。用户将能够给该数据和音乐注释以‘活动标签’（我当时在跑步，或者正在做作业）其将开辟其它形式的关于唤醒、效价、音乐、其它娱乐体验和活动之间的关系分析。

[0300] 该应用将使个体能够在他们社交网络中搜索具有类似的情绪，或从事类似的活动的人，例如‘在我的网络中寻找某个想要说话的人’或者情绪低落但热衷于谈话的人。这能够通过情绪板指示出来或在 Facebook 和其它社交网络上扩大状态信息。

[0301] 随着大量的用户，参加（在匿名表达规则和同意分享的影响下）的人表达他们自动产生的情绪，该数据可指示在群体和更大的社区之间的全部唤醒状态。

[0302] 该应用将被扩展，以提供图形化的网络地图，其示出社会团体之间的情绪的模式和群集，创造一种‘社会情感’景象，在社会环境中，其针对从事他们自己的个体活动或者聚集在一起的活动群体，例如，在社交聚会中，或听音乐会，或跳舞。

[0303] 这与社交网络分析的早期实例相反，其受源自语言和语义分析的数据挖掘和模式匹配的限制，并因此限制了他们的精确度。X-系统将通过捕捉真实的先天神经生理状态信息，产生个体和群体唤醒更真实和准确的解释。

[0304] 该应用通过将X-系统的用户连接到网站的cookies，也将用于优化网站，这样，如果我正在浏览网站并同意使用X-系统读取我的先天状态信息，cookies将产生网站或特定网页对情绪影响的分析。这将使网站设计者尝试各种文本、影片、音乐和画面显示，布局和体验以及获得关于用户的情绪反应的即时反馈。

[0305] 该信息之后将用于与广告和营销索引进行匹配，这样，对网站体验的反应可以与品牌价值和创造特定产品和服务目标的期望的情绪或要求进行匹配。因此，例如，反馈机制可用于匹配关于特定汽车的广告的情绪反应。

[0306] X-系统的核心算法的这种扩展创造了一个新的连通形式，其在深层次运行，超越了文化的界限、语言的情绪表达，选择性地将其连接到当前活动中，包括音乐的选择、其它的娱乐活动以及其它的活动。

[0307] 该无意识的、前语言水平的唤醒、情感和效价的连通，开辟了一个新的社交网络和医疗诊断的范例。在医疗机构中，例如，监测病人的'状态'信息将提供自知力而不是利用传统的诊断技术。X-系统可以与各种传统的医疗、护理和诊断装置和应用程序集成，以创造病人病情和情绪状态的更全面的描绘。

[0308] X-系统关于先天唤醒、效价和音乐选择的核心数据，通过标准界面传递给广泛使用的社交网络，例如，Facebook和Twitter，以及直接传递给局域网络中的智能手机。

[0309] X-系统将被嵌入到智能手机和其它装置中，以与各种软件、固件和芯片硬件的多种组合的方式。X-系统API将使专业应用程序开发人员能够创造各种工具和技术，以利用'状态'信息流，创造反馈和监测服务。

[0310] 有很多将数据和界面传递给社交网络和智能手机的协议和系统。X-系统的该应用是独一无二的，因为，其能够以新数据扩展这些系统，否则将无法使用。X-系统扩展到目标的先天唤醒和效价的连通，其具有在本地、广域和社交网络中的个人或群体中指示此时的音乐、其它娱乐活动或自发的活动的随附数据。

[0311] X-系统也能够分享唤醒值给搜索引擎，该唤醒值与用户和搜索引擎，例如Google®，之间的互动相关。该搜索引擎随后可以使用这些值优化该搜索引擎的搜索和广告选择。

[0312] X-系统也能够分享唤醒值，其与用户在具有网站优化系统的网站中浏览特定网站或网页相关，以便网站优化系统能够使用这些值来优化网站和/或特定的网页(内容、布局、声音等)。

[0313] K.扩展/增强的时机

[0314] 对产品的改进和扩展的主要方向如下：

[0315] ●识别对音乐的情绪反应是受记忆或对歌词的刺激或歌曲或一段音乐的其它方面的反应，而不是生物学——由过滤掉预期的生理反应开发。

[0316] ●传感器的发展和配件，例如，新一代的微型脑电图描记器(EEG)、大脑扫描传感器。一种可能的方法是在耳机和手套中包括传感器(测试任何上述的参数，例如，脉搏、皮肤电导等)。

[0317] ●高级的音乐搜索、导航和发现系统。

[0318] ●高级的音乐搜索、导航和发现系统，包括推广、订购、选择和控制界面。

[0319] ●专业的医学应用。

[0320] ●分析音乐，以确定先天的情绪反应 ;和

[0321] 从早期采用者中捕获和分析传感器数据，以微调整唤醒水平。

[0322] 有两个用于细化分析功能的进一步的策略。第一个是通过系统的大规模使用。其计划招募一百名志愿者分五个阶段来测试该系统。他们在听音乐过程中的生理数据，包括心率和皮肤电导的读数，该数据将与自动分类数据和手动分类的结果进行比较，作为在自动分析程序中识别优势和劣势的手段，无论是在数据的捕获还是值的组合中。第二个用于细化的策略是通过机器学习，利用线性回归和 / 或神经网络方法。训练阶段将在五个测试阶段的每一个之后。该方法将具有这些价值，即，审阅现存值和它们的组合，并建立所了解的信息和程序的发展资源。其可能不能够显著地细化自动分类。如果这被证明是事实的话，机器学习程序和统计分析将被用于产生必要的细化。此外，在自动分类系统中的劣势能够通过收集和分析特定音轨对用户产生的影响的真实的测试来修正。本领域的技术人员将认识到，可以使用人工智能 (AI) 和基于探试规则的方法，以及迭代自动化和测试方法论。

[0323] X- 也能够用于在零售环境、和 / 或在线社区中，通过播放合适的音乐来创造和调节‘情绪’。个体能够通过网站界面连接，以产生常见的反应 / 读数。

[0324] 类似地，X- 系统能够用在理解和匹配对品牌的情绪反应中——基本上是利用 X- 系统作为一种工具，用来诊断并且之后分享对品牌的情绪反应，这是通过准确地将这些品牌与针对目标受众的合适的音乐关联实现的。X- 系统能够用于判断不同社交群体对品牌音乐的反应。

[0325] 利用轮询和类似的人群传感技术，X- 系统也能够作为在群体环境中的动态群体诱导工具使用，以选择加剧唤醒的音乐，例如在运动或娱乐活动中，以及用来在公共环境中减少群体的紧张和挫败，例如交通工具、医院和政府大楼中。

[0326] L.X- 系统的优势

[0327] 该技术预期会具有广泛的社会、生理和生物优势，在减少压力、治疗失眠中，在集中和学习的优化中，在提高创造性思维中，以及在促进最佳的运动模式中，无论是普通的人群或是针对优秀运动员的培训制度，还用在提高活动的竞争力中。

[0328] X-系统可以适用在针对特定医疗条件的治疗方法中。有大量的文献提供音乐药物和音乐疗法的功效的证据，其作为对于以下状况的治疗的补充支持，例如，慢性疼痛、痴呆、帕金森氏病、抑郁症、创伤后紧张症和失语症，以及在缓和剂、手术后、中风后护理。可能的优势包括手术后减少卧床休息时间和减少药物使用时间。

[0329] 作为一个实例，简希望能够更加集中在手头的工作上，因此她滑动腕带式无线传感器，触摸她的 iPhone 上的“集中”符号，并且当她继续工作时听。该系统将监测她的思想和身体状态，并播放合适的音乐来保持适当的集中水平。

[0330] 应该注意的是，此外，X- 系统的自动分类算法具有相当大的潜在的市场价值，因为其“独一无二”，不依赖于传感器技术，能够针对音乐成流系统提供一种“情绪的”导引能力。

[0331] 本发明可有利地用于根据音乐的神经生理效果选择和分类音乐，包括但不限于，根据音乐的神经生理影响，进行音乐的排序 / 定序、使用、宣传、订购和销售。本发明也可以

有利地用于将这种分类连接到其它常用的分类方案中。

[0332] 该系统其它潜在的用途可以是从音乐数据库中选择合适的音乐片段用于在电影中的配音，其期望观众的特定情绪。其也能够用在视觉艺术中，其期望观众的特定情绪。因此，这些应用程序将是视觉应用程序或视听应用程序，而不仅是音频应用程序。

[0333] 相关的产品和服务将从这样两个方面产生，以产生关于市场中的发展趋势的市场情报，即，涉及音乐分析的产品和服务，以确定先天的情绪反应，以及将产生来自早期采用者的传感器数据的捕获和分析以微调整唤醒水平，以产生关于在未来市场中的趋势的情报。实例中可包括服务于电子游戏产业以促进声道选择从而提高互动游戏技术的情绪体验，或者作为对音乐作曲家寻求引起对所推荐的乐曲的整体或部分产生特定的反应的辅助。

[0334] 注释

[0335] 应当理解的是，上面引用的设置仅仅是针对本发明的原理的应用的说明。可以在不背离本发明的精神和范围的情况下做出更多的修改和其它的安排。虽然本发明已在附图中示出，并且结合目前认为是本发明最实用和优选的实施例，详细具体地在以上充分说明，对本领域的技术人员显而易见地是，在不背离如本文所列的本发明的原理和概念的情况下可以做出更多的修改。例如，在本文中给出的数学方程仅是特定的和非限制性的实例。

[0336] 附录 1

[0337] X- 系统技术要点 :组件概述

[0338] 从根本上，X- 系统包含 3 个组件，其中的两个是软件，另一个是硬件。

[0339] 一个软件（“音乐分析器”）用在脱机的（不与系统的实时操作直接连接）模式中，以分析候选音乐文件，以及建立他们的兴奋 / 情感影响的评估。

[0340] 第二个软件部分是回放组件。其负责实际播放音乐文件并且也负责接收来自于传感器硬件的数据，并利用它来升级它的内部模型，该内部模型确定之后要播放的音轨。

[0341] 最后，硬件组件包含多个传感器，其收集来自于本地环境、主要是实际用户的实时数据。

[0342] 详细说明

[0343] 音乐分析

[0344] 音乐分析子系统的分析方面已经在别处详细描述，在这里不再赘述。本部分仅描述集成方面。如上所述，预期其主要在脱机时以非互交方式运行。其将针对于一批音乐输入周期地运行，其结果是产生描述音轨的某些属性的一组值。这些值也能够结合用于产生音轨的单一的‘兴奋’图，其由回放系统使用。单独存储组件的优势是，当数据被收集用于调谐系统时，在不需要重新分析整个音轨的情况下，兴奋值能够随着不同的系数被重新计算，这大大地减少了开销。

[0345] 所有的分析输出将被存储在数据库中，并以多个参数被索引，至少包括音轨和艺术家标识符，并且其中一些形成声学特征，其是相对宽容的编码差异和背景噪声。

[0346] 这些索引将在用户‘输入’他们的音乐收藏时使用。如果有任何曲目已经在数据库中存在，它们的值将不需要被重新计算。

[0347] 反馈程序将是一种选择系统（opt — in system），在该系统中，用户同意提供关于他们的匿名信息供系统使用，以改进该系统。

[0348] 可以使用自动化功能，例如，唤醒中的归一化变化、建议的音轨的重放 / 跳过、和合计的传感器数据。也可以使用以喜欢 / 不喜欢的应答以及偶尔的随机问卷调查的形式的显示反馈。

[0349] 使用反馈以诱导系统参数可以以总体和每个用户为基础。大规模的数据挖掘、模式识别、机器学习系统将用于改进音乐的情感 / 唤醒评估。

[0350] 分析组件将作为互联网访问服务运行，结合一些音乐成流服务以提供音频，或者纯粹作为运行用户的个人音乐收藏的控制系统。

[0351] 在可以使用快速和可靠的互联网服务的地方，程序的重要组分能够被卸载到托管的 X- 系统服务中。这允许比典型的终端装置更强大的程序，并且也保护了分析器的 IP。

[0352] 附加的用途

[0353] 除了‘唤醒调整’的主要目标——促进放松或兴奋——还存在音乐分析的其它可能的用途。其可以用于将额外的维度加入到音乐发现和导引中，这通过观察大量的短音乐样本对用户产生的影响，并且之后建议类似特征的音轨或艺术家给用户。如果该系统已经被一些人在任何合理的时间使用，并且具有适应良好的个人模式，将不需要初始步骤。类似于“像艺术家 / 专辑 X 的音乐”的导航，也有可能是基于音轨分析中所确定的特征。

[0354] 回放和决定

[0355] 回放组件处理两个任务。基于传感器输入来控制音乐回放和运行实时的唤醒分析 / 诱导模型。该组件可负责实际播放音乐，或者可以是在现有媒体播放器——例如，iTunes/Windows 媒体播放器等——的基础上的控制层。唤醒分析模型将基于 X- 系统 INRM 模型，使用来自于作为起始点的音乐分析组件的预先计算的值。用户将选择期望的输出，以及传感器将用于测量接近每个音轨输出的进度。显式的超控装置将允许用户一次手动跳过特定的音轨，或者永久地将其计入黑名单中，以保证其永远不再被选择。除了它们的效果之外，这些超控装置将给决定模型提供回馈。

[0356] 该组件的功能将在一定程度上取决于其正在运行的环境。在相关的低功率的装置上，例如，手机和便携式音乐播放器，其可能以较小的精确度、较小的计算集约模式运行，或者如果可能的话，卸载一些程序到远程服务中。

[0357] 针对笔记本电脑 / 台式机 / 平板电脑的应用程序，可使用更复杂的模型。针对小生境的使用，其可与模拟器或视频回放组件相结合运行，以增强诱导的效果。

[0358] 很多用户很可能希望在多个不同的主机上使用该系统，例如，他们的手机和平板电脑上。该播放器需要一些在系统之间同步和分享这些模型数据的方法。这可能最好通过（或基于）一些类似于苹果 iCloud 或 GooglegDrive 的互联网服务实现。这也将提供用于显示数据给分析系统从而用于模拟 / 训练的渠道。

[0359] 附加的用途、注释

[0360] 给予足够的训练，也许有可能开发一个 X- 系统版本，其能够某种程度上在没有传感器反馈的情况下运行。其效果很可能低于机械化装置，但是对于用户可以有足够的价值，在于避免传感器购买、维修和穿着不便的复杂化。如果这被证明是不可能或不可取的，将很可能通过没有直接连接用户的传感器获得一些反馈，例如，在他们的口袋里携带的手机的加速计，或者在同时指示他们的位置的 GPS。

[0361] 传感器硬件

[0362] 目前,系统的传感部分使用两个不同的传感器。一个是脉搏血氧仪,其用于监测心率,另一个是皮肤电导传感器,其测量皮肤的电传导性(电阻的倒数)。

[0363] 脉搏血氧仪

[0364] 脉搏血氧仪在通过血流中(氧)血红蛋白的光吸收波长原理的基础上运行。通过比较红光和红外的波长吸收值,可以确定含氧血的相对比例,产生'血氧饱和度'(spO2)图。以相对高的频率追踪该值,其允许指示心跳引起的脉搏的突然变化的检测,因此,能够确定心跳的速度。尽管在医疗环境中很有用,血氧不发生显著地变化或在对X-系统有用的时间尺度上不发生显著地变化或,并且在只有心率数据被收集。

[0365] 当前系统使用COTS(商品化的产品和技术)传感器,燕牌3150腕带式氧传感器无线脉搏血氧仪。该装置使用软橡胶指尖夹来覆盖光线发射器/检测器,其是典型的传感器类型。可选择地,可使用传感器轻轻夹紧耳朵的耳垂,以及身体的其他方面的装置。该装置使用蓝牙(具有标准和通用的SPP-串行端口协议)用于数据传输。

[0366] 在该传感器的将来的实施例中,很可能使用比指尖更便利和更少侵入性的传感器位置。通过利用直接传递吸收(即,光直接穿过相对较薄的身体部分,例如手指或耳垂),传感器的可靠性和准确性被大大地改善,但是可以以反射模式运行的装置确实存在,允许其几乎可以放在任何位置,虽然区域要求是血管密度高的位置,并且优选相对靠近皮肤表层的位置。一个好的位置——即,非常适合X-系统的目标的位置——将是表带的一部分,具有在手腕内侧的传感器,其中,表带扣位于典型的表带上。

[0367] 皮肤电导

[0368] 皮肤电导,有不同的术语,EDA(电子皮肤活性)、GRS(皮肤电阻抗),或仅是皮肤电阻/电导,其是皮肤带电的能力的测量。基于显而易见的安全因素,电流必须保持在非常低的值,并且需严格控制。针对个人和他们的本地环境的基线皮肤电导率取决于多个因素,但是在短时间尺度上,主要的影响因素是汗液。汗液,本质上是富含电解质的水,是一种良好的导体,并且它的存在降低了皮肤的有效电阻。顺便说一句,电导率(由西门子/欧姆测得(Siemens/mhos))被定义为电阻的倒数(以欧姆表示)。按照惯例,在描述这些系统的时候使用电导,虽然其转换为电阻是微不足道的。

[0369] 出汗受多种因素的影响,但是我们更感兴趣其与副交感神经系统的关系。唤醒增强与出汗增加有很大的关联性,从而使皮肤电导增加。这种效果是比较快的,可数秒完成。身体汗腺密度最高的区域——手和脚的工作表面——是最有效的传感位置,但是其它位置也是可能的,结果各不相同。手腕和前臂外侧已经显示出可以提供合适的结果[可供参考]。

[0370] 测量皮肤电导可以多种方式实现。目前的传感器使用简单的分压器,其具有高精度电阻作为接脚,并且两个皮肤接触器应用于用户以作为另外的接脚。为了测量,中央节点也与缓冲的ADC(数模转换器)以用于测量。

[0371] 也存在其它的设计,并且已经利用惠斯登电桥——特殊的电路结构,其允许高精度微分测量——做了一些原型产品,以改进精度和噪声抑制。

[0372] 该参数的一个重要的方面是,值可以变化几个数量级。干性皮肤,在寒冷的、干燥的环境中,可以具有微西门子(micro — Siemen)(兆欧姆)范围的电导,而且极爱出汗的皮肤可以降到数百的毫西门子(milli-Siemen)(1 — 1000欧姆)。在这么大的范围内的精确

测量，对于传感器的设计提出了重大的挑战。

[0373] 现有的传感器，如上所述的，使用相对简单的分压器。其通过集成在传感器微控制器（MCU）中的模拟－数字转换器（ADC），在 50Hz 左右采样。

[0374] 目前使用的特定的 MCU 是德州仪器（Texas Instruments）MSP430F2774。除了 ADC 之外，该装置包含集成的可编程的增益放大器（PGA），其用于放大来自于 1x 至 16x 的信号。这提供了从 4 位到现有的 10 位 ADC 的精度的有效增加。前述的分放大器是另一个用在跟随器（增益）模式的集成的运算放大器，其运行以缓冲信号，并且提供高阻抗负载给分压器，以保证读数不会因为流经采样子系统的大电流而失真。

[0375] ADC 输入是在近似 50Hz 的范围采样。如果测试的值落入靠近其全部测量范围的上端或下端的两个区域之一，PGA 的前置放大器的增益将被调整，以将其提高来接近测量范围的中心。紧随本次调整之后（在放大器所需的短的稳定期之后）将采用另一个样本。一个滞后的方法在每个区域的边缘实现，以使两个放大器增益水平之间的‘翻滚’的可能性以及与及时采集的值的干扰被反复地减到最少。此外，相对高的采样速率（50Hz）与约 2Hz 的传输速率相比，留有足够大的空间给放大器调节。高采样率的读数利用具有 10Hz 的保险装置的简单的低通（FIR）滤波器平均化。

[0376] 一旦第二个样本完成，落入这些边缘区域并导致放大的变化的样本将被删除。软件信号灯用在固件中，以保证通信子系统当其使用或包含不可靠的数据的时候不能访问样本缓冲器。

[0377] 如果读数落入缓冲区域但是前置放大器可能已经设置到最大值或最小值，该读数将被存储和传输，但标记有指示潜在的饱和 / 消波错误的标记。

[0378] MCU 也与无线电模块连接，其用于与 USB 基站连通。无线通信在同样不受管制的频带运行，例如 WiFi 和蓝牙，在 2.4GHz。然而，它们具有非常低的功率和数据速率，并被设计成与这些附近的其它装置共存。

[0379] 更高级别的无线电通信的运用是利用在传感器装置和基站上的 SimpliciTI（简单网络协议）专有网络协议的稍加修改的版本。其允许多个传感器在彼此的范围内运行，同时保证数据被正确的基站接收。基站是利用第二个 MSP430 实现，这次有 USB 接口，并且其使用标准的 USB 串行装置驱动器，该驱动器由几乎所有的主机装置和操作系统支持。分层的网络协议之上是 X- 系统传感器协议，其存在主要是为了传感器读数的方便传输、提供调试输出，以及为了节约功率而允许传感器的选择性启用 / 禁用。传感器升级的频率也可以调节。

[0380] 该传感器是采用电池供电的，可以通过 USB 在原位充电。这允许完全无线操作，并且使可能出现在外部电源线的噪音最小化。

[0381] 注释

[0382] 上述部分描述了现有的实施例，但是还有许多计划但尚未展开的附加特征。这些包括升级当前的传感模式，以及与传感器附加类型的结合。

[0383] 升级包括：

[0384] ●心率：

[0385] ○适合手腕式传感的反射红外脉搏血氧计。

[0386] ○针对心脏波形和节律分析的高频体积描记的采样，超过了简单的‘心率’值。

[0387] ●皮肤电导：

[0388] ○基于皮肤电导传感器的惠斯登电桥，其具有分散或集成的精密仪表放大器。

[0389] ○更复杂的数字滤波阶段。

[0390] ○使用附属于／靠近皮肤接触的同步加速器来将因为接触式移动物品的读数标记为可疑。

[0391] 附加的特征包括：

[0392] ● EEG 类型的传感器或针对脑波活动的‘帽子’

[0393] ●肌电图仪肌张力／触发率

[0394] ●针对高分辨率的心脏波形的多点 ECG

[0395] ●呼吸深度／速度

[0396] ●眼球追踪／注视／眨眼分析

[0397] 未来的数据源尚不可行，但是其将有利于该系统，包括：应激激素（例如皮质醇）血浆浓度、神经触发速率、区域的大脑活动。

[0398] 在传感器的发展中，将要克服的主要障碍是便利。如果针对大众市场，很少的用户会忍受繁琐的连接线或他们的手或感官的障碍，例如，与治疗或医疗市场相比。将传感器合并成单一的程序包中，例如手表或耳机这样的用品，将是理想的方式。其它的可行性包括将柔性电路集成于服装或鞋类中。

[0399] 传感器包应该有能力与尽可能多的主机装置具备可行的互用性。这可包括智能手机、功能手机、平板电脑、便携式音乐播放器、手提电脑、家庭音响、和车载音频。最常用的接口可能是 WiFi 或蓝牙，虽然支持所述主机的范围明显不同。

[0400] 附录 2

[0401] 模拟人类神经生理反应

[0402] 下列文件，其通过引用并入本文，在模拟人类的神经生理反应方面提供了信息。

[0403] Aragon D，Farris C，Byers JF，竖琴音乐在血管中和胸外科手术病人中的效果 (The effects of harp music in vascular and thoracic surgical patients)，保健与医学替代疗法 (Alternative Therapies in Health and Medicine)，20029 － 10，;8(5)：52-4,56-60；Baumgartner T，Lutz K，Schmidt CF，Jancke L，音乐的情感力量：音乐如何增强情感图像的感觉 (The emotional power of music：how music enhances the feeling of affective pictures)，脑研究 (Brain Research)，20062；1075(1)：151 － 64；Bernardi L，Porta C，Sleight P，在音乐家和非音乐家中不同类型的音乐引起的心血管，脑血管呼吸的变化：消除噪音的重要性 (Cardiovascular，cerebrovascular and respiratory changes induced by different types of music in musicians and non-musicians：the importance of silence)，心脏（英国心脏学会）(Heart (British Cardiac Society))，20064；92(4)：445 － 52；Blood AJ，Zatorre RJ，对与涉及奖励和情感的大脑区域中的活动关联的音乐的强烈愉悦的反应 (Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion)，美国国家科学院院刊 (Proceedings of the National Academy of Sciences)，2001925；98(20)：11818-23；Brown S，Martinez MJ，Parsons LM，被动式音乐聆听将自发地连接边缘和旁边缘系统 (Passive music listening spontaneously engages limbic and paralimbic systems),神经报告(Neuroreport),2004915;15(13):2033-7;Brugge JF,在听觉皮层中的组织模式(Patterns of organisation in auditory cortex),美国声学学会杂志(Journal of the Acoustical Society of America),78(1/2)1985353-359;Byers JF,Smyth KA,音乐干预对心脏手术病人的噪音滋扰、心率和血压的影响(Effect ofa musical intervention on noise annoyance,heart rate,and blood pressure in cardiac surgery patients),美国危症护理杂志(American Journal ofCritical Care),19975;6(3):183-91;Cardigan ME,Caruso NA,Haldeman SM,McNamara ME,Noyes DA,Spadafora MA,Carroll DL,音乐对卧床休息的心血管病人的影响(The effects ofmusic on cardiac patients on bed rest),心血管病护理进展(Progress in Cardiovascular Nursing),2001冬;16(1):5-13;Decety J,Chaminade T,感觉同情的神经管理(Neural correlates offeeling sympathy),神经心理学(Neuropsychologia),412003127-138;Evers S,Suhr B,在短时间内音乐感知过程中血清素变化但荷尔蒙未变化(Changes of the neurotransmitter serotonin but not of hormones during short time music perception),欧洲神经病学文献与临床神经科学(European Archives of Psychiatry and Clinical Neuroscience),2000;250(3):144-7;Formisano E,Kim DS,Di Salle F,van de Moortele PF,Ugurbil K,Goebel R,在人类初级听觉皮层中的镜面对称音质映射(Mirror-symmetric tonotopic maps in human primary auditory cortex),神经元(Neuron),40(4)2003859-869;Gallese V,移情作用的根源。共享歧管假说和主体间性的神经基础(The roots of empathy.The shared manifold hypothesis and the neural basis of intersubjectivity),精神病理学(Psychopathology),362003171-180;Gerra G,Zaimović A,Franchini D,Palladino M,Giucastro G,Reali N,Maestri D,Caccavari R,Delsignore R,Brambilla F,健康志愿者对‘高技术音乐’的神经内分泌反应:与人格特质和情绪状态的关系(Neuroendocrine responses of healthy volunteers to ′ techno-music ′ :relationships with personality traits and emotional state),国际心理生理学杂志(International Journal of Psychophysiology),19981;28(1):99-111;Grape C,Sandgren M,Hansson LO,Ericson M,Theorell T,唱歌能够促进健康?:专业和业余歌手歌唱课程中的实证研究(Does singing promote well-being?:An empirical study of professional and amateur singers during a singing lesson),综合生理和行为的科学(Integrative Physiological and Behavioral Science),20031-3;38(1):65-74;Griffiths TD,Buchel C,Frackowiak RS,Patterson RD,人类大脑对声音中的时间结构的分析(Analysis of temporal structure in sound by the human brain),自然神经科学(Nature Neuroscience),1(5)1998422-427;Hebert S,Beland R,Dionne-Fournelle O,Crete M,Lupien SJ,生理应激对玩视频游戏的反应:内置音乐的贡献(Physiological stress response to video-game playing:the contribution of built-in music),生命科学(Life Sciences),200541;76(20):2371-80;Holstege G,Bandler R,Saper CB(编),情感运动系统(The emotional motor system),大脑研究进展(Progress in Brain Research)107,爱思唯尔(Elsevier),阿姆斯特丹1996;Iwanaga M,心率和偏爱的音乐节奏之间的关系(Relationship between heart rate and preference for tempo of music)，知觉和运动技能(Perceptual and Motor Skills),199510；81(2)：435-40；Iwanaga M，KobayashiA，Kawasaki C，反复接触音乐的心率变异性(Heart rate variability with repetitive exposure to music)，生物心理学(Biological Psychology),20059；70(1)：61－6；Iwanaga M,Tsukamoto M,兴奋和镇静的音乐对主观和生理放松的影响(Effects of excitative and sedative music on subj ective and physiological relaxation),知觉和运动技能(Perceptual and Motor Skills),19978；85(1)：287-96；Jeannerod M,视觉和行动线索有助于自我-他人的区分(Visual and action cuescontribute to the self-other distinction)，自然神经科学(Nature Neuroscience),7(3)2004421－422；Knight WE Rickard NS,轻松的音乐可防止诱导增加在健康的男性和女性中的主观焦虑、收缩压和心率(Relaxing music prevents stress-induced increases in subj ective anxiety，systolic blood pressure and heart rate in healthy males and females)，音乐治疗杂志(Journal of Music Therapy),2001冬；38(4)：254-72；Koelsch S，Fritz T，V Cramon DY，Muller K，Friederici AD,音乐情感的调查：功能磁共振成像研究(Investigating emotion with music：an fMRI study),人类大脑成像(Human Brain Mapping),20063；27(3)：239-50；Kumar AM，Tims F，Cruess DG，Mintzer MJ，Ironson G，Loewenstein D，Cattan R，Fernandez JB,Eisdorfer C，Kumar M,音乐疗法增加了患阿耳滋海默氏病病人的血清褪黑激素水平(Music therapy increases serum melatonin levels in patients with Alzheimer′s disease),保健与医学替代疗法(Alternative Therapies in Health and Medicine),199911；5(6)：49-57；Lee DN,耦合T介子的指导运动(Guiding movement by coupling taus),生态心理学(Ecological Psychology),1998：10(3-4)：221－250；Lee OK,Chung YF,Chan MF,Chan WM,Music and its effect on the音乐及其对接受机械通风的病人的生理反应和焦虑水平的影响：试点研究(Music and its effect on the physiological responses and anxiety levels of patients receiving mechanical ventilation：a pilot study),临床护理杂志(Journal of Clinical Nursing),20055；14(5)：609-20；Li L，Korngut LM，Frost BJ，Beninger RJ,下神经丘的病变后的前脉冲抑制：前脉冲强度函数(Prepulse inhibition following lesions of the inferior colliculus：prepulse intensity functions),生理学与行为(Physiology and Behavior),19988；65(1)：133－9；McAdams S，Winsberg S，Donnadieu S，De Soete G，KrimphoffJ，合成的音乐音质的知觉比例：通用维度、特异性、和潜在的主题类型(Perceptual scaling of synthesised musical timbres：common dimensions，specidicities，and latent subj ect classes),心理学研究(Psychological Research),581995177－192；Meloni EG，Davis M,蜗背侧核有助于大鼠的听觉惊跳反射的高强度组元(The dorsal cochlear nucleus contributes to a high intensity component of the acoustic startle reflex in rats)，听觉研究(Hearing Research),19985；119(1－2)：69-80；MenonV，Levitin DJ,Smith BK,Lembke A,Krasnow RD,Glazer D,Glover GH,McAdams S,在谐波声音中的音色变化的神经关联(Neural correlates oftimbre change in harmonic sounds),神经成像(Neuroimage),17(4)20021742－1754；Miluk-Kolasa B，Obminski Z，Stupnicki R，Golec L，音乐治疗对患者在手术前紧张而分泌的唾液皮质醇的影响

(Effects of music treatment on salivary cortisol in patients exposed to pre-surgical stress)，实验与临床内分泌学(Experimental and Clinical Endocrinolog),1994;102(2):118-20;Mok E，Wong KY，音乐对病人的焦虑的影响(Effects of music on patient anxiety),美国手术期护理杂志(AORN Journal),20032;77(2):396-7,401－6,409－10;Molnar-Szakacs I,Overy K,音乐和镜像神经元:从运动到‘电子’运动(Music and mirror neurons:from motion to′ e′ motion),社会认知情感神经科学(Social Cognitive Affective Neuroscience1),2006235-241;Nilsson U，Unosson M,Rawal N,病人在手术后接触平静的音乐可减少压力和镇痛:一项随机对照试验(Stress reduction and analgesia in patients exposed to calming music postoperatively:a randomized controlled trial)，欧洲麻醉学杂志(European Journal ofAnaesthesiology),20052;22(2):96－102;Osborne N1,在冲突和冲突后区域中的儿童音乐(Music for children in zones of conflict and post-conflictin.),交际音乐性编(Communicative Musicality ed)，Malloch和Trevarthen,牛津大学出版社(OUP)2009;Osborne N2，接近音乐节奏的时间生物学(Towards a chronobiology ofmusical rhythmin)，交际音乐性编(Communicative Musicality ed)，Malloch和Trevarthen，牛津大学出版社(OUP)2009;Osborne N3，神经科学和现实世界的实践...(Neuroscience and real world practice...)纽约客学院记事(Annals of the New York Academy of Sciences)2012(出版中);Overy K，Molnar-Szakacs I,在一起的时间:音乐体验和镜像神经元系统(Being together in time:musical experience and the mirror neuron system),音乐知觉(Music Perception),262009489-504;Pachetti C，Aglieri R，Mancini F，Martignoni E，Nappi G,主动音乐治疗和帕金森氏病(Active music therapy and Parkinson′ s disease:methods)，功能性神经学(Functional Neurology),19981－3;13(1):57－67;Panksepp J，Trevarthen C,情感神经科学(The neuroscience of emotion in musicin),交际音乐性编(Communicative Musicality),牛津大学出版社(OUP)2009;Pantev C,Hoke M,Lehnertz K,Lutkenhoner B,Anogianakis G，Wittkowski W,瞬时听觉诱发磁场脑电图揭示的人类听觉皮层的音质组织(Tonotopic organisation of the human auditory cortex revealed by transient auditory-evoked magnetic fields)，临床神经生理学(Electroencephalographic Clinical Neurophysiology),69(2)1988160-170;Patterson RD,Uppenkamp S,Johnsrude IS，Griffiths TD,听觉皮层中的时间间距和旋律信息的处理(The processing of temporal pitch and melody information in the auditory cortes)神经元(Neuron)36(4)2002767-776;Penhune VB，Zatorre RJ，Feindel WH,在颞叶清除——包括黑索氏回——的病人中听觉皮层在节奏模式保留中的作用的研究(The role of auditory cortex in retention of rhythmic patterns as studied in patients with temporal lobe removals including Heschl′ s gyrus),神经心理学(Neuropsychologia).19993;37(3):315-31;Peretz I，听大脑:对音乐情绪的生物反应(Listen to the brain:a biological perspective on musical emotionsin);Juslin P,Sloboda J(编),音乐和情感:理论和研究(Music and Emotion:Theory and Research)牛津大学出版社,伦敦2001;Peretz I，Zatorre RJ(编)，音乐的认知神经科学(The cognitive neuroscience ofmusic) 牛津大学出版社,伦敦 2003 ;Peretz I,Kolinsky R,音乐鉴别中节律之间的可分性边界 :以神经心理学的角度 (Boundaries of separability between rhythm in music discrimination :a neuropsychological perspective),实验心理学季刊 (The Quarterly Journal ofExperimental Psychology)19935 ;46(2) :301 - 25 ;Reinhardt U,癌症患者在放松疗法中的心率和音乐节律的研究 (Investigations into synchronisation of heart rate and musical rhythm in relaxation therapy in patients with cancer pain),(德语)Forschende Komplementarmedizin19996 ;6(3) :135-41 ;Rencanzone GH,Schreiner CE,Merzenich MM,成年猫头鹰猴在辨别训练之后的初级听觉皮层的频率表示中可塑性 (Plasticity in the frequency representations of primary auditory cortex following discrimination training in adult owl monkeys) 神经系统科学 (Neuroscience)13(1)199387 - 103 ;Rizzolati G,Fogassi I,Gallese V. 以理解和模仿动作为基础的神经生理结构 (Neuro-physiological mechanisms underlying the understanding and imitation of action) 神经系统科学自然评论 (Nature Reviews Neuroscience),2,2001661 - 670 ;Schneider N,Schedlowski M,Schurmeyer TH,Becker H,通过音乐使经历脑血管造影术的病人减轻压力 (Stress reduction through music in patients undergoing cerebral angiography) 神经放射学 (Neuroradiology)20016 ;43(6) :472-6 ;Stefano GB,Zhu W,Cadet P,Salamon E,Mantione KJ,音乐本质地改变收听者中的快速麻醉和细胞因子过程 (Music alters constitutively expressed opiate and cytokine processes in listeners), 医学科学评议 (Medical Science Monitor)20046 ;10(6) :MS18-27 ;Sutoo D,Akiyama K,音乐改进多巴胺能的神经传递 :基于音乐对血压调节的实例 (Music improves dopaminergic neurotransmission :demonstration based on the effect of music on blood pressure regulation) 脑研究 (Brain Research)200486 ;1016(2) :255 - 62 ;Talavage TM,Sereno MIO,Melcher JR,Ledden PJ,Rosen BR,Dale AM,由频率灵敏度的级数揭示的人类听觉皮层中的音质组织 (Tonotopic organisation in human auditory cortex revealed by progressions of frequency sensitivity) 神经生理学杂志 (Journal of Neurophysiology)91(3)20041282 - 1296 ;Trevarthen C,音乐性和内在的音乐脉冲 :来自于人类的神经生物学和婴儿沟通的证据 (Musicality and the Intrinsic Motive Pulse :Evidence from human psychobiology and infant communication),音乐科学特刊 :人际交往的节奏、故事和起源 (Special Issue of Musicae Scientiae :Rhythm,Narrative and Origins of Human Communication)1999 :157-213 ;Trevarthen C,Malloch SN,健康之舞 :定义音乐治疗效果 (The Dance of Wellbeing :Defining the Musical Therapeutic Effect) 北欧音乐治疗杂志 (Nordic Journal of Music Therapy)2000 ;9(2) :65 - 126 ;Turner R,Ioannides AA,大脑、音乐和音乐性 :神经影像的推论 (Brain,music and musicality :inferences from neuroimaging) 交际音乐性 (Communicative Musicality) 牛津大学出版社 2009 ;Uedo N,Ishikawa H,Morimoto K,Ishihara R,Narahara H,Akedo I,Ioka T,KajiI,Fukuda S,在结肠镜检查中降低唾液皮质醇水平的音乐疗法 (Reduction in salivary cortisol level by music therapy during colonoscopic examination),肝-胃肠病学 (Hepato-gastroenterology)20043-4 ;51(56) :451 - 3 ;Updike PA,Charles

DM,音乐治疗方案：等待整形手术的术前患者对录制的音乐节目的生理和情绪反应(Music Rx：physiological and emotional responses to taped music programs of preoperative patients awaiting plastic surgery),整形外科杂志(Annals of Plastic Surgery).19877；19(1)：29-33；Urakawa K,Yokoyama K,音乐能够增强由心率变异性评估的运动诱发的交感神经优势(Music can enhance exercise-induced sympathetic dominancy assessed by heart rate variability),

[0404] 日本东北实验医学杂志(The Tohoku Journal ofExperimental Medicine)20057；206(3)：213－8；VanderArk SD,Ely D,生物和音乐专业的大学生对不同偏好值的音乐刺激的皮质醇、生物化学、和皮肤电反应(Cortisol,biochemical,and galvanic skin responses to music stimuli of different preference values by college students in biology and music),知觉和运动技能(Perceptual and Motor Skills).19938；77(1)：227-34；Warren JD,Uppenkamp S,Patterson RD,Griffiths TD,在人类大脑中分离音高浓度和音高高度(Separating pitch chroma and pitch height in the human brain),美国国家科学院院刊(Proceedings of the National Academy of Sciences)美国,100(17)200310038－10042；Wieser HG,Mazzola G,音乐和音和不和谐度：它们是通过左右海马单独地辨别吗？(Musical consonances and dissonances：are they distinguished independently by the right and left hippocampi?)神经生理学(Neuropsychologia)1986；24(6)：805－12；Yamamoto T,Ohkuwa T,Itoh H,Kitoh M,Terasawa J,Tsuda T,Kitagawa S,Sato Y,运动前听慢和快节奏的音乐对超大的循环性能和选择的代谢变量的影响(Effects ofpre-exercise listening to slow and fast rhythm music on supramaximal cycle performance and selected metabolic variables)生物化学与生理学文献(Archives of Physiology and Biochemistry)20037；111(3)：211-4；Zatorre RJ,Peretz I(编)音乐的生物学基础(The biological foundations ofmusic),纽约科学院(New York Academy of Sciences)2001；Zatorre R,J,Evans AC,Meyer E,以针对音高的旋律感知和记忆为基础的神经机制(Neural mechanisms underlying melodic perception and memory for pitch),神经科学杂志(Journal ofNeuroscience)14(4)19941908－1919。

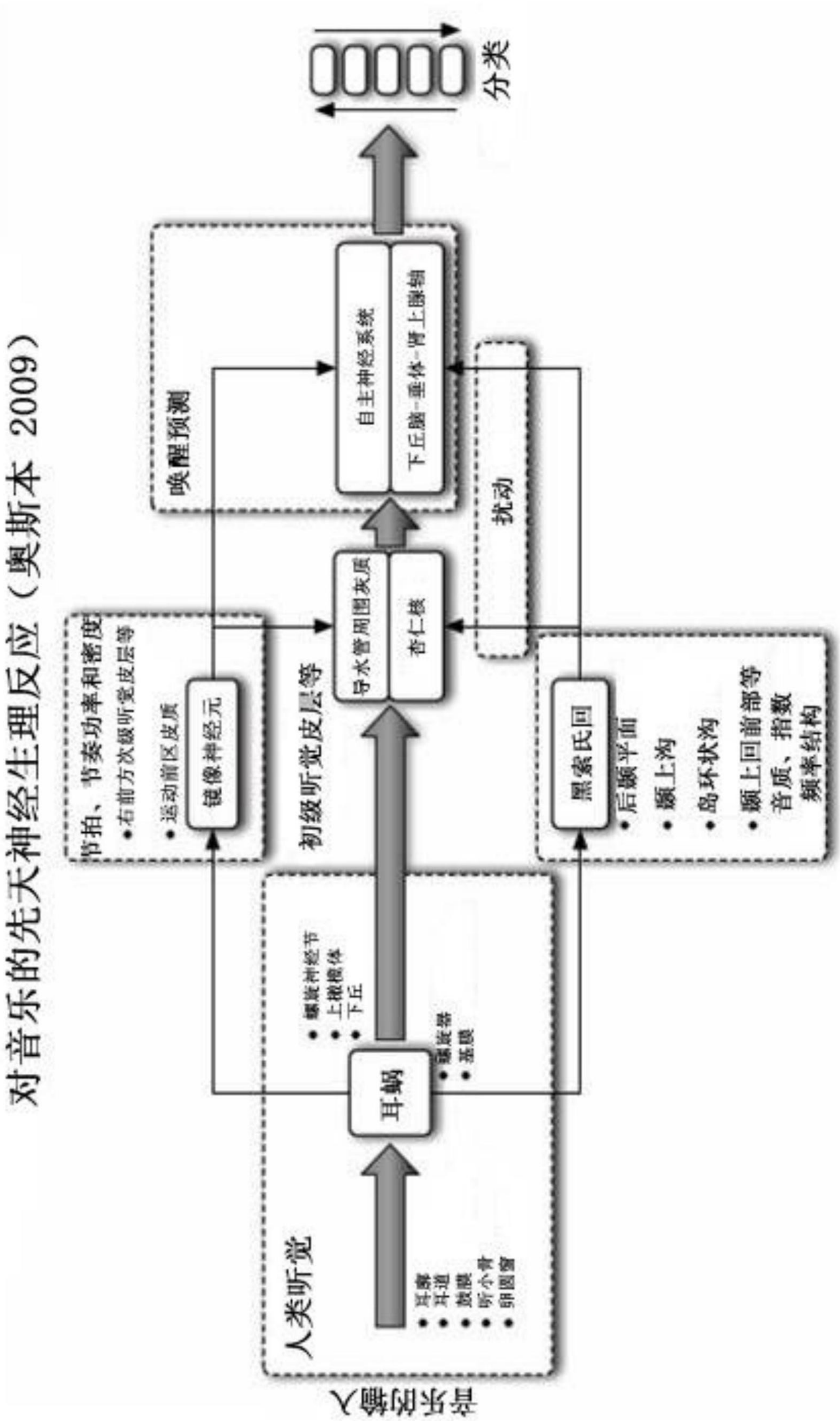

对音乐的先天神经生理反应（奥斯本 2009）
音乐的输入
人类听觉
耳廓
耳道
鼓膜
听小骨
卵圆窗
耳蜗
螺旋神经节
上橄榄体
下丘
螺旋器
基膜
初级听觉皮层等
导水管周围灰质
杏仁核
节拍、节奏功率和密度
右前方次级听觉皮层等
运动前区皮质
镜像神经元
黑索氏回
后颞平面
颞上沟
岛环状沟
颞上回前部等
音质、指数
频率结构
扰动
唤醒预测
自主神经系统
下丘脑-垂体-肾上腺轴
分类

图 1

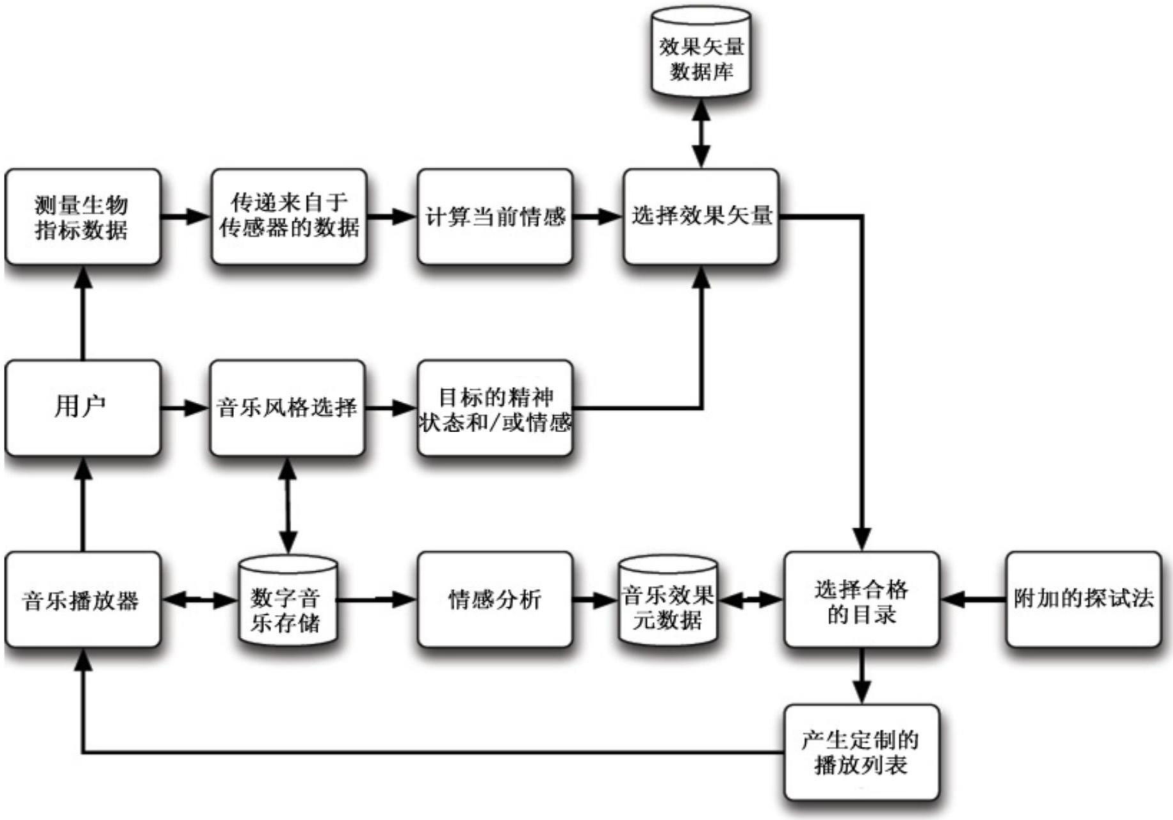

效果矢量数据库
测量生物指标数据
传递来自于传感器的数据
计算当前情感
选择效果矢量
用户
音乐风格选择
目标的精神状态和/或情感
音乐播放器
数字音乐存储
情感分析
音乐效果元数据
选择合格的目录
附加的探试法
产生定制的播放列表

图 2

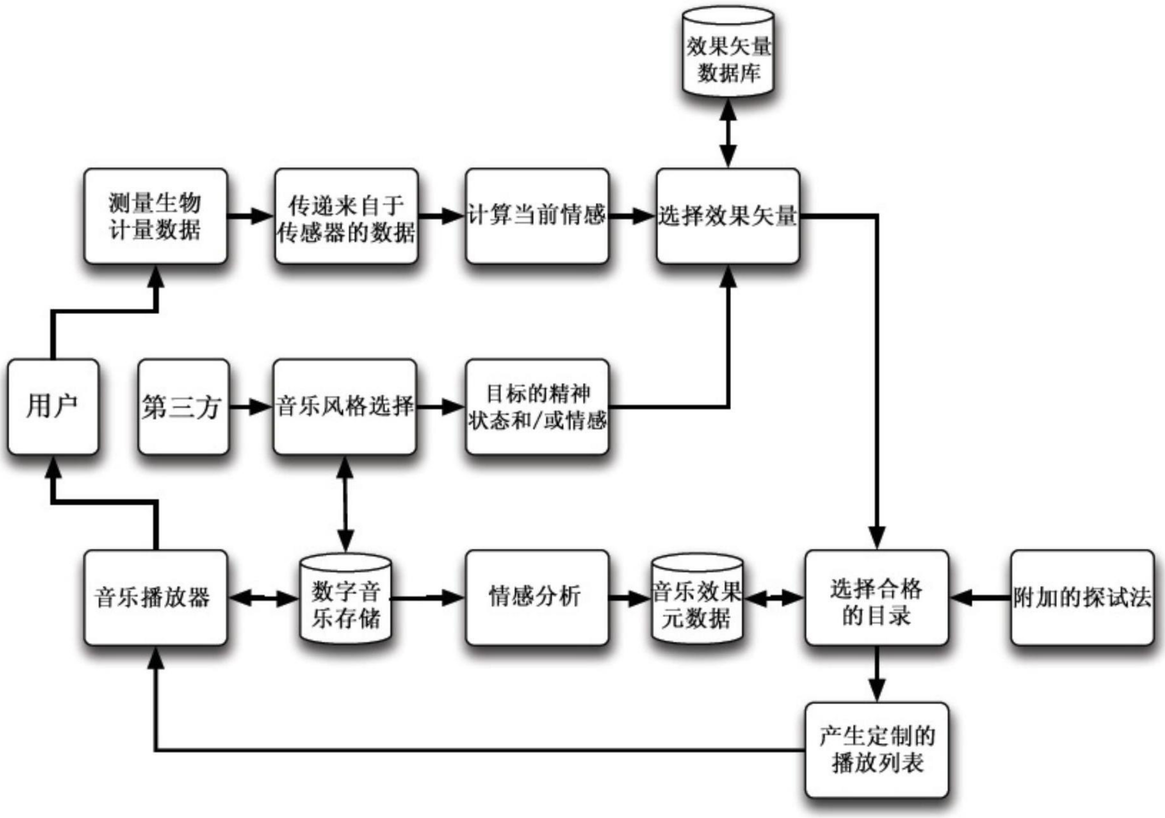

效果矢量数据库
测量生物计量数据
传递来自于传感器的数据
计算当前情感
选择效果矢量
用户
第三方
音乐风格选择
目标的精神状态和/或情感
音乐播放器
数字音乐存储
情感分析
音乐效果元数据
选择合格的目录
附加的探试法
产生定制的播放列表

图 3

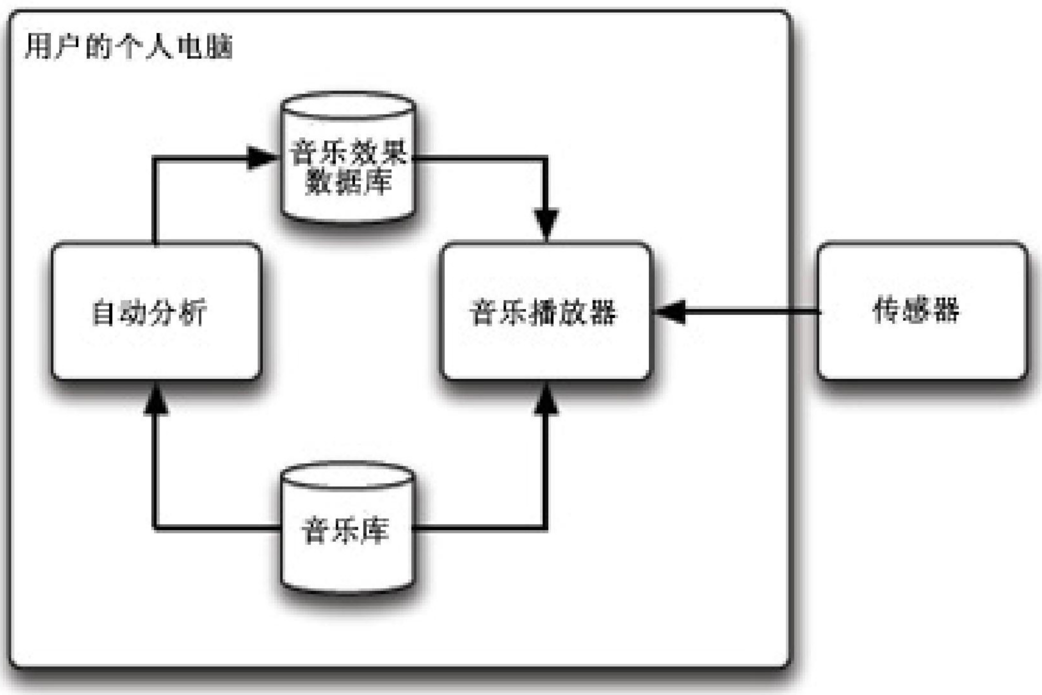

用户的个人电脑
音乐效果数据库
自动分析
音乐播放器
传感器
音乐库

图 4

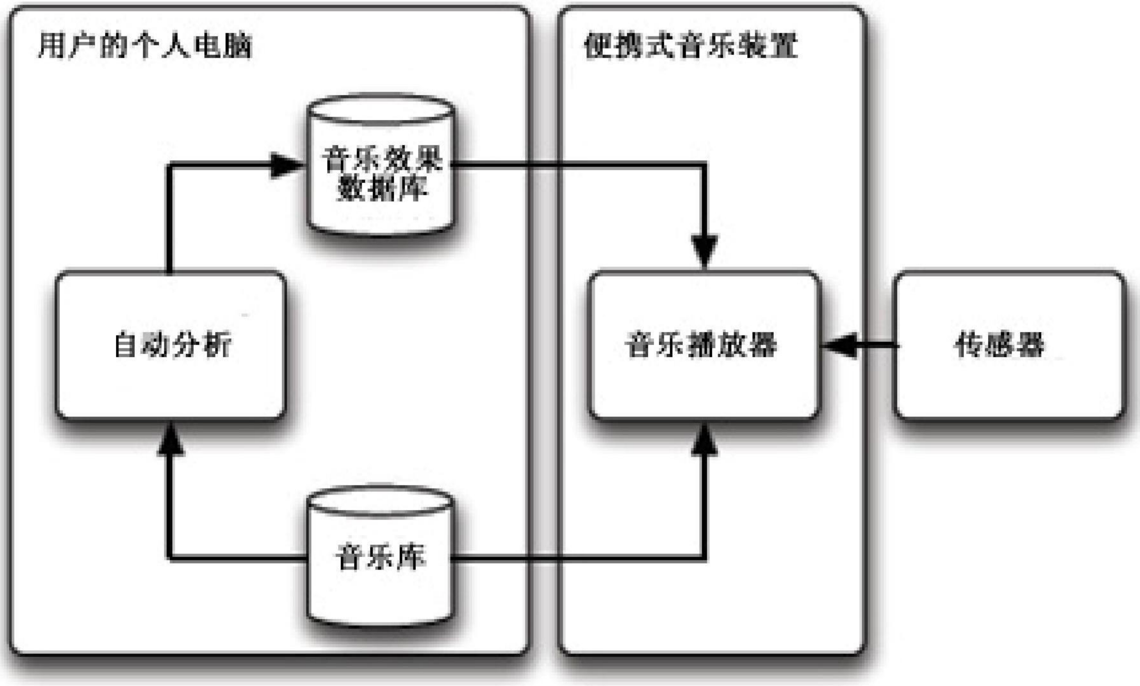

用户的个人电脑
便携式音乐装置
音乐效果数据库
自动分析
音乐播放器
传感器
音乐库

图 5

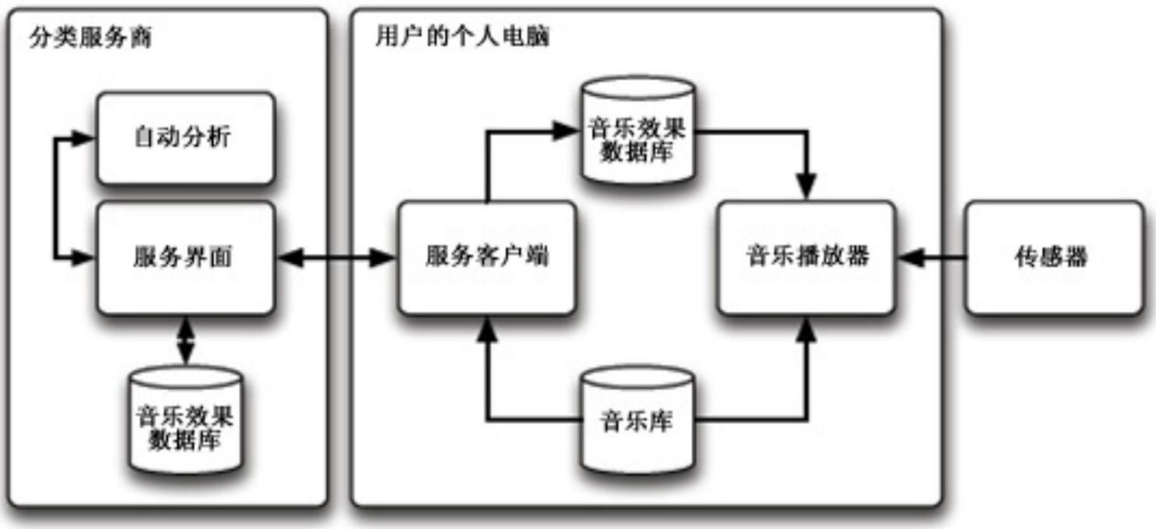

分类服务商
用户的个人电脑
自动分析
服务界面
音乐效果数据库
服务客户端
音乐效果数据库
音乐播放器
传感器
音乐库

图 6

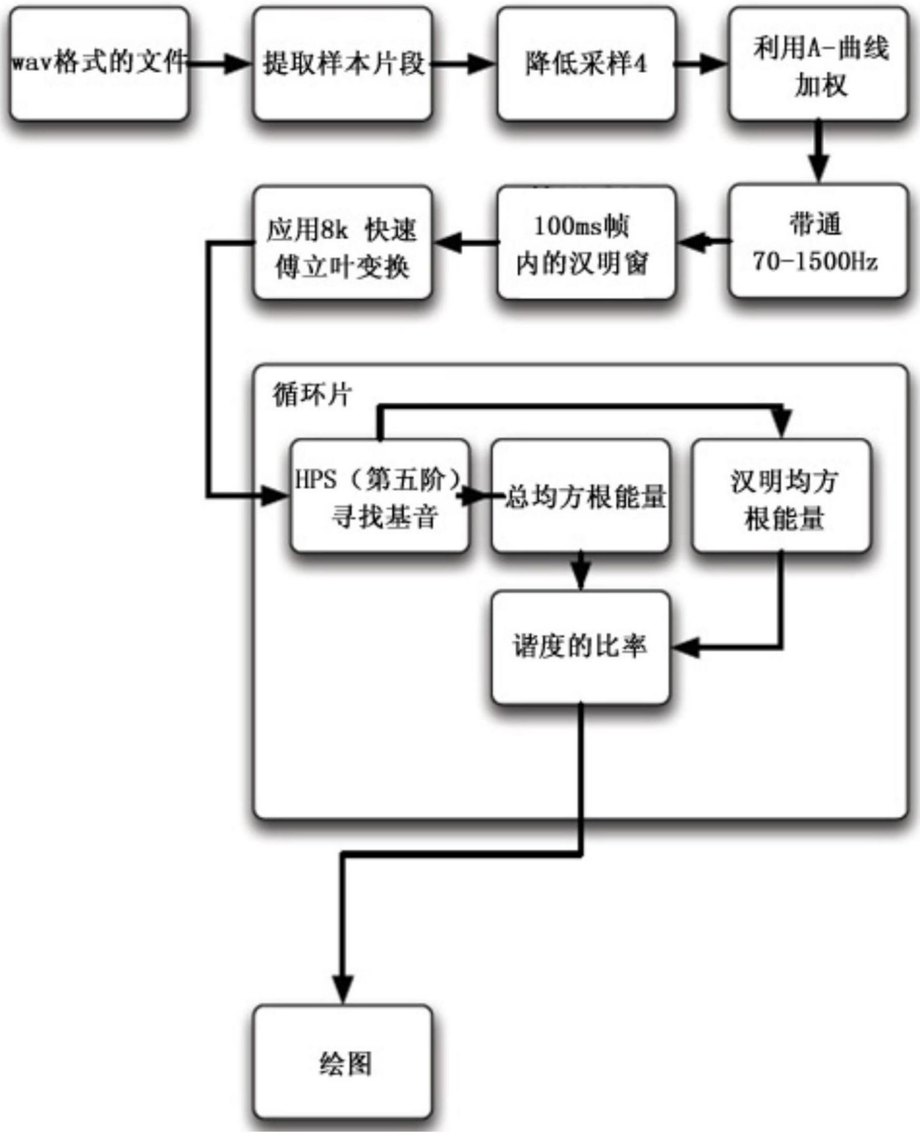

wav格式的文件
提取样本片段
降低采样4
利用A-曲线
加权
带通
70-1500Hz
100ms帧
内的汉明窗
应用8k 快速
傅立叶变换
循环片
HPS（第五阶）
寻找基音
总均方根能量
汉明均方
根能量
谐度的比率
绘图

图 7A

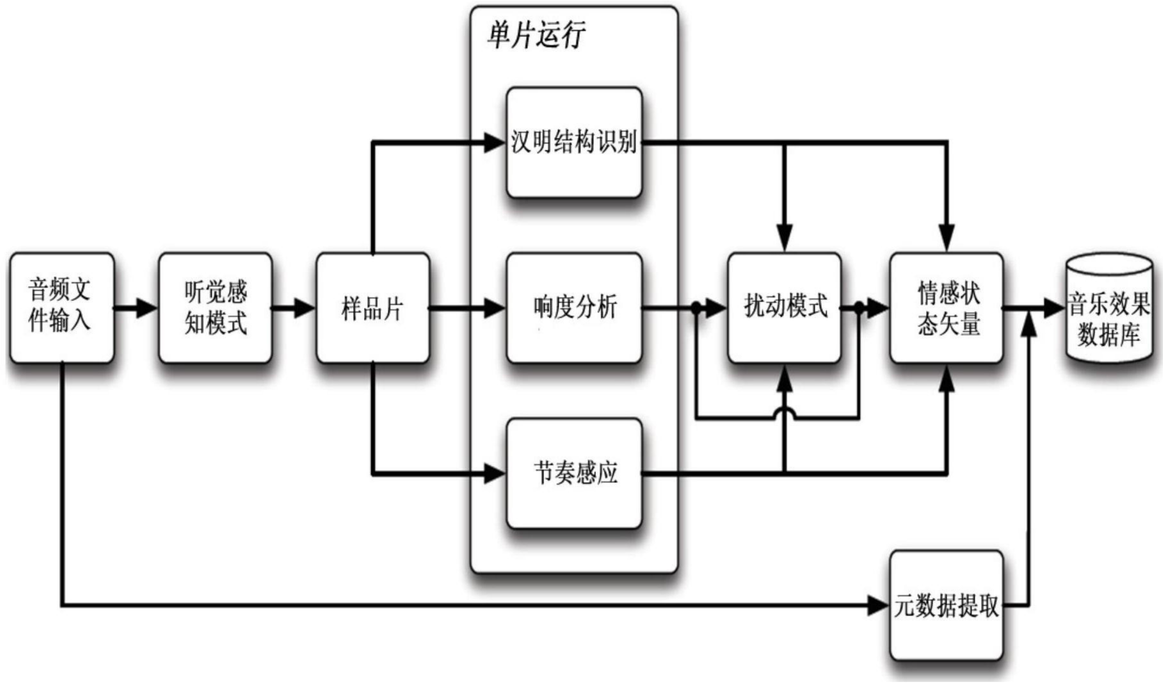

单片运行
汉明结构识别
音频文件输入
听觉感知模式
样品片
响度分析
扰动模式
情感状态矢量
音乐效果数据库
节奏感应
元数据提取

图 7B

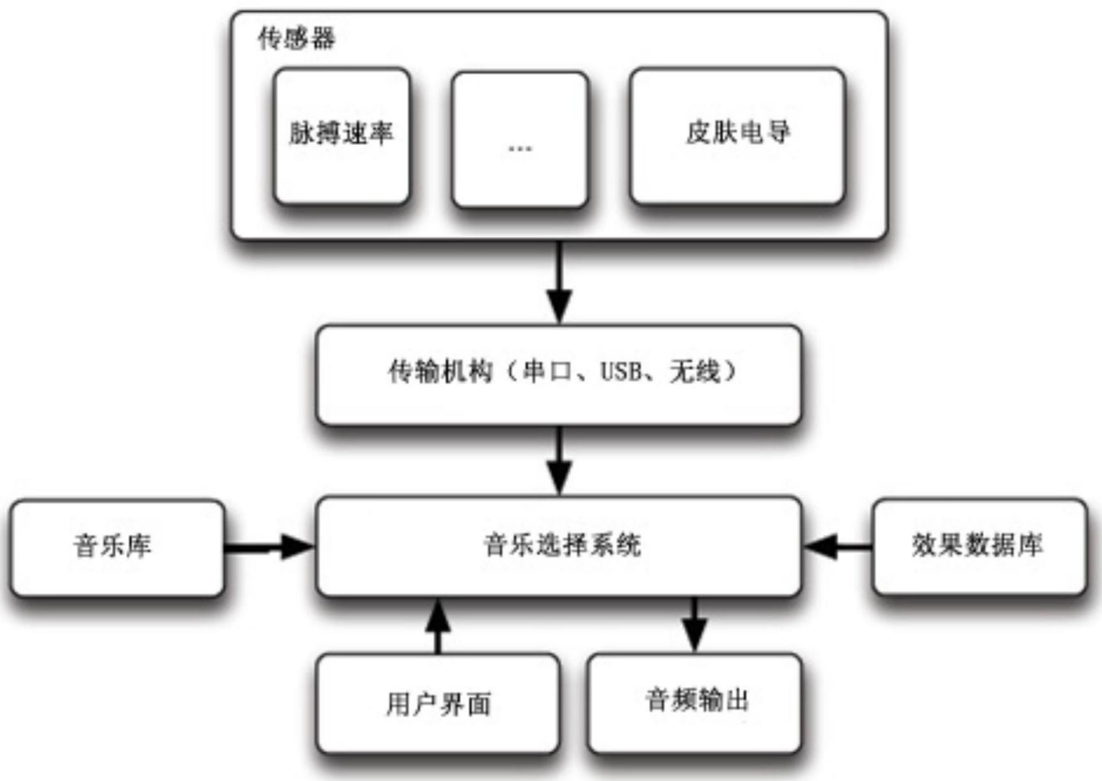

传感器
脉搏速率
...
皮肤电导
传输机构（串口、USB、无线）
音乐库
音乐选择系统
效果数据库
用户界面
音频输出

图 8

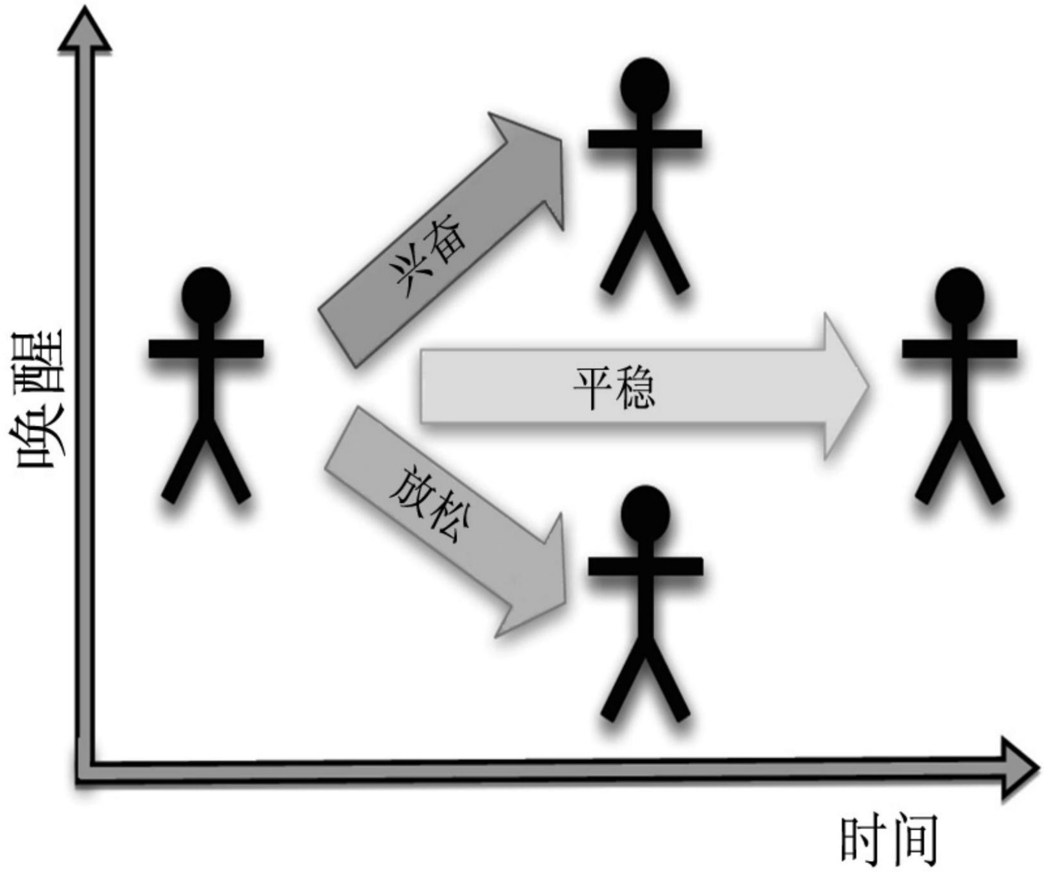


图 9

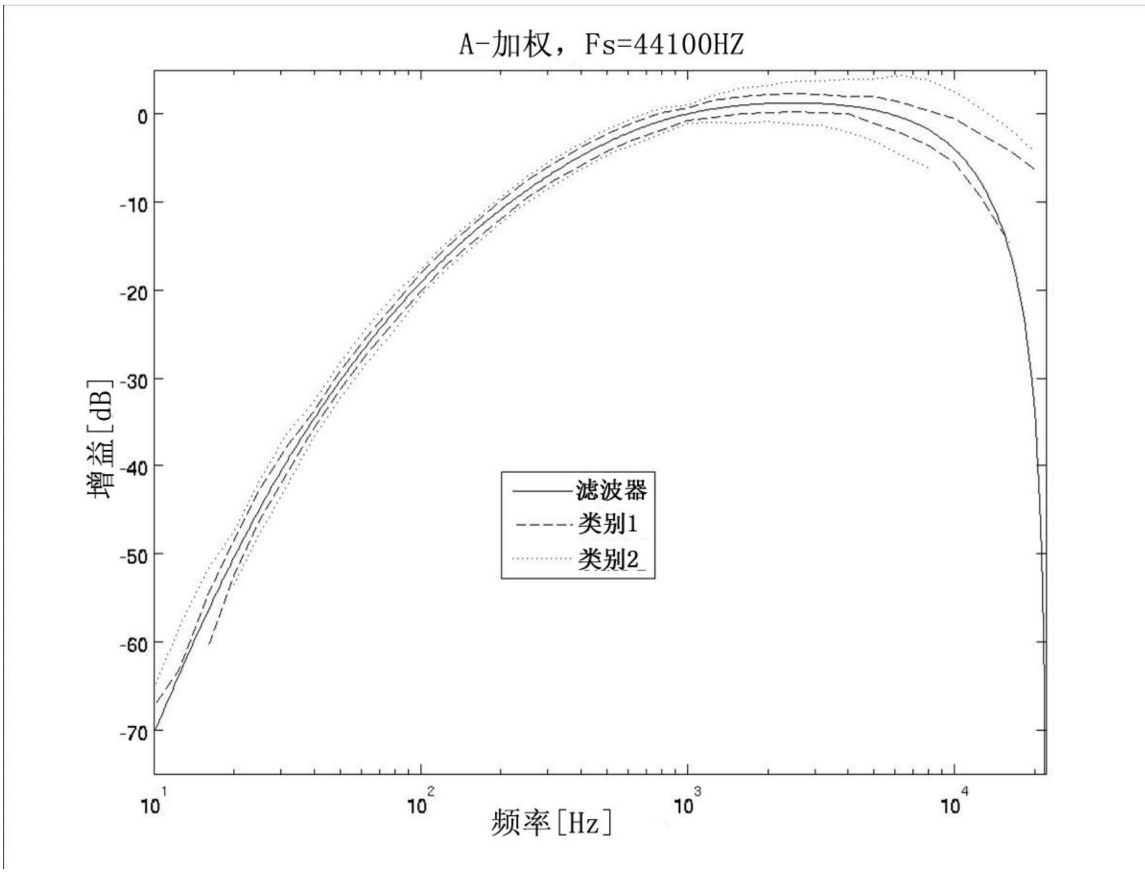

A-加权，Fs=44100HZ
增益[dB]
0
-10
-20
-30
-40
-50
-60
-70
10^1
10^2
10^3
10^4
频率[Hz]
滤波器
类别1
类别2

图 10

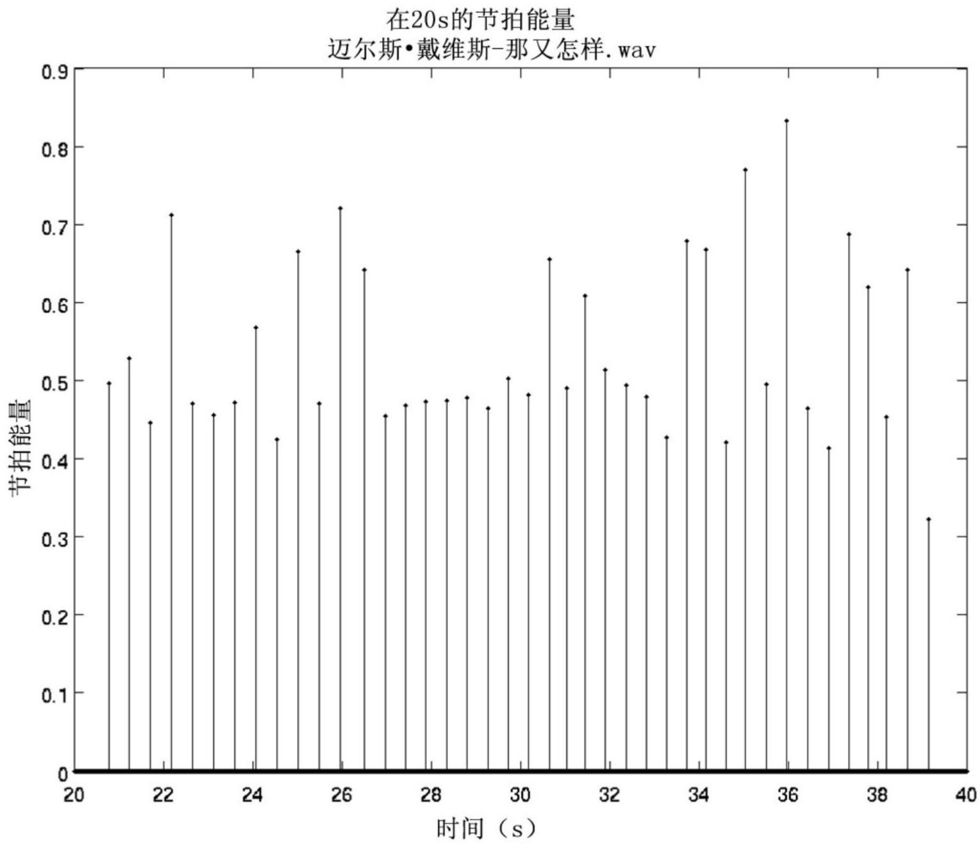
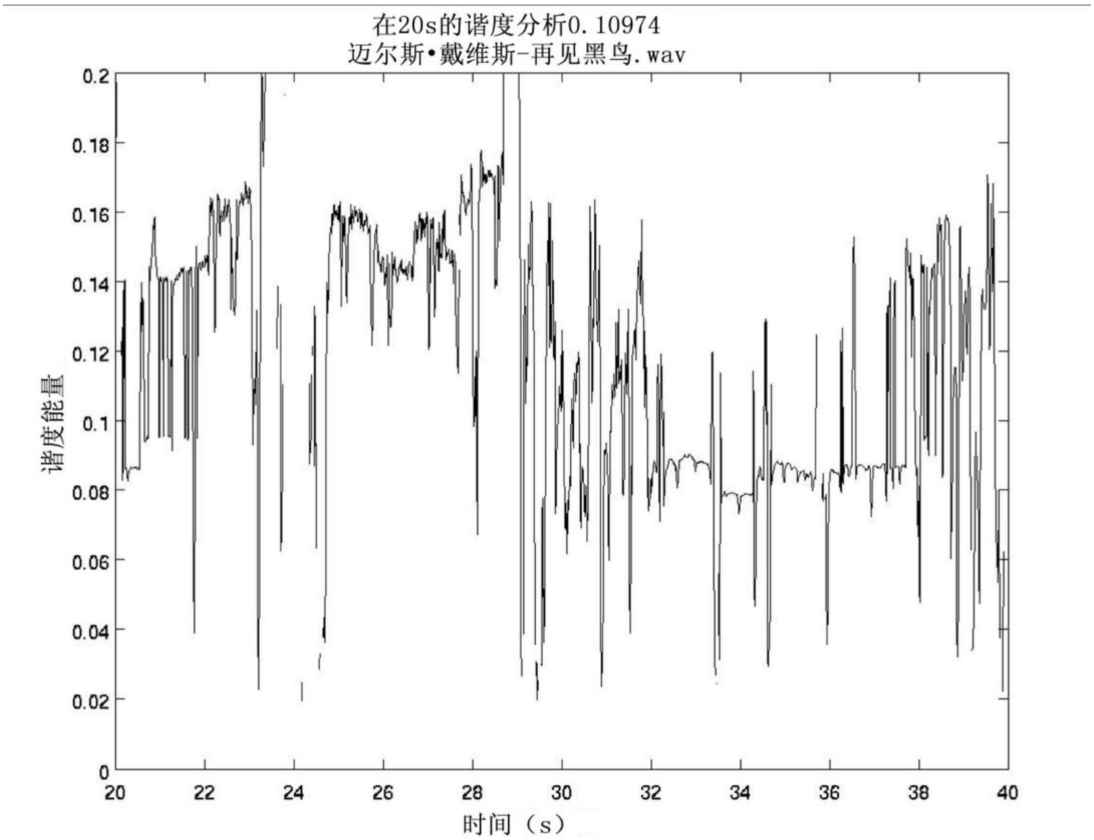
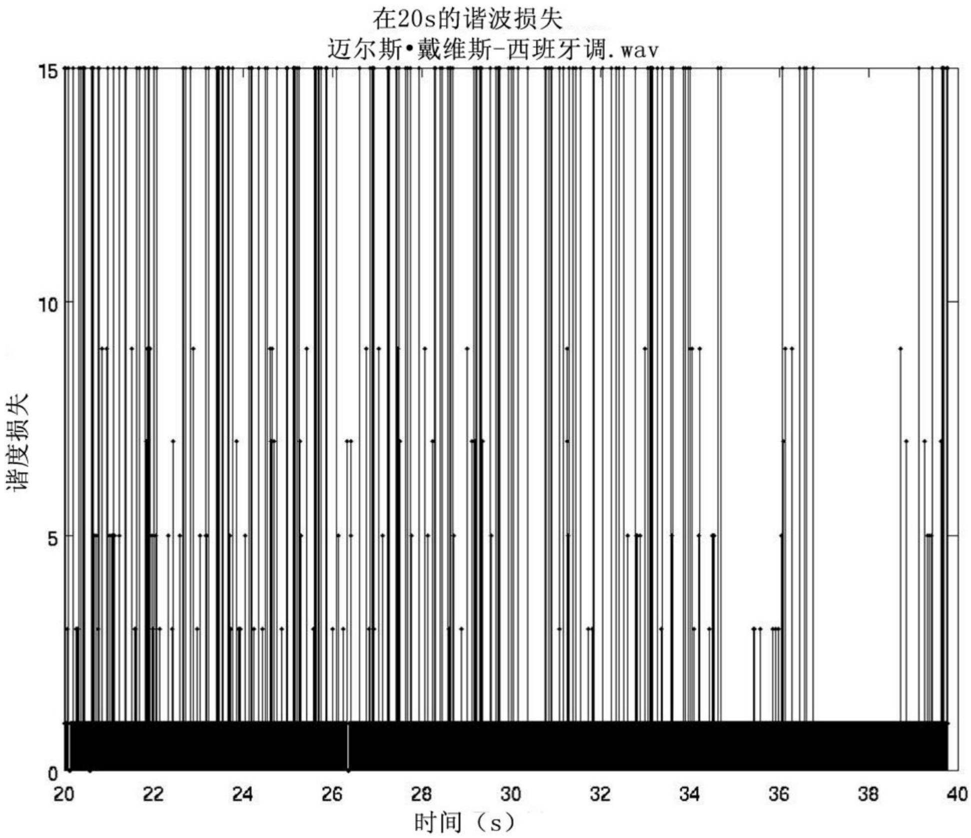

在20s的节拍能量
迈尔斯•戴维斯-那又怎样.wav
节拍能量
时间（s）
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
20
22
24
26
28
30
32
34
36
38
40

图 11



在20s的谱度分析0.10974
迈尔斯•戴维斯-再见黑鸟.wav
谱度能量
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
20
22
24
26
28
30
32
34
36
38
40
时间（s）

图 12



在20s的谐波损失
迈尔斯•戴维斯-西班牙调.wav
谐度损失
15
10
5
0
20
22
24
26
28
30
32
34
36
38
40
时间（s）

图 13

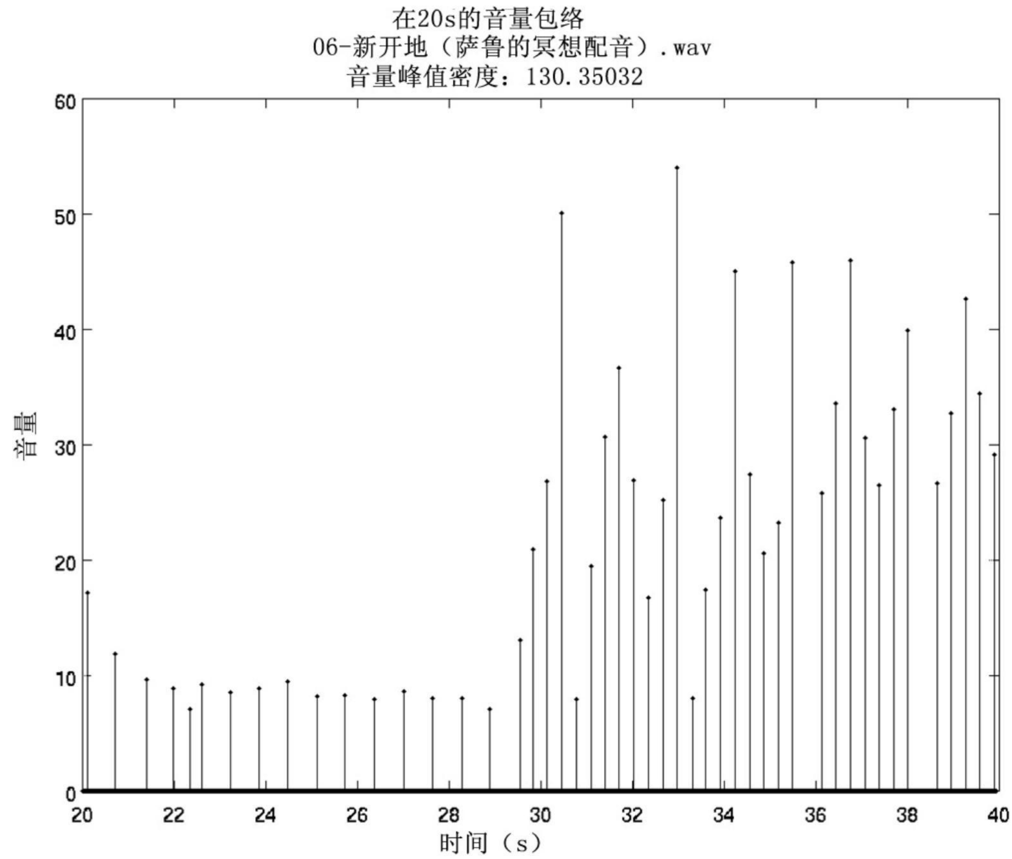
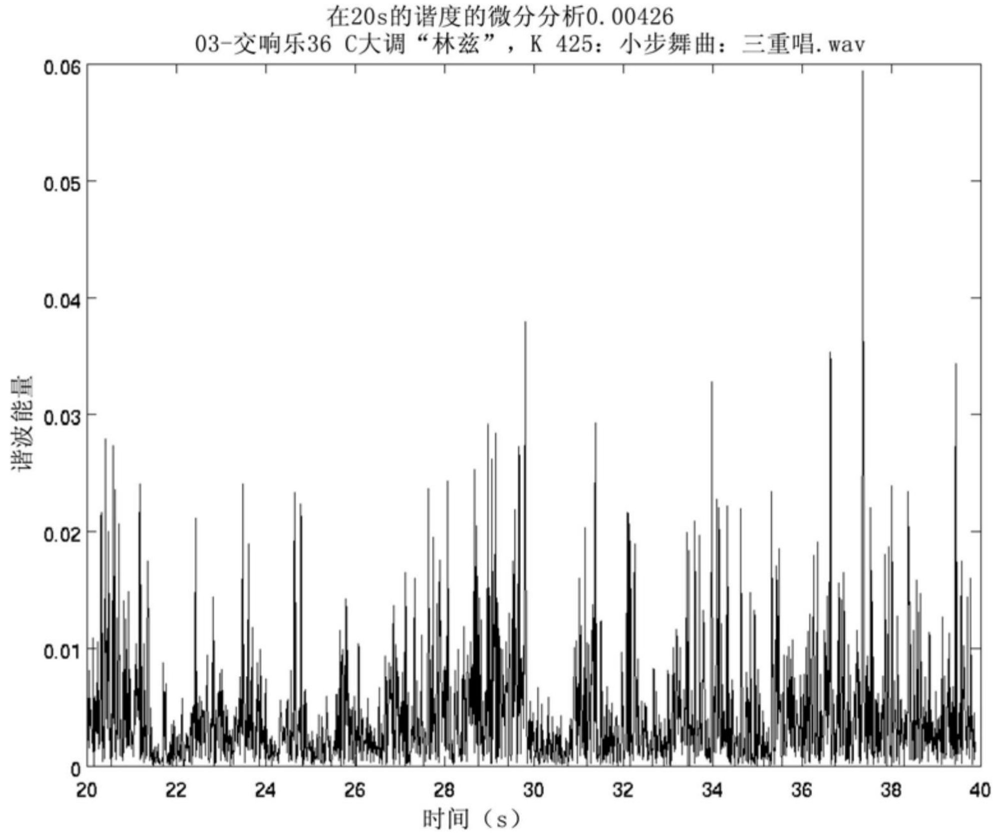

在20s的音量包络
06-新开地（萨鲁的冥想配音）.wav
音量峰值密度：130.35032
音量
时间（s）
0
10
20
30
40
50
60
20
22
24
26
28
30
32
34
36
38
40

图 14



在20s的谐度的微分分析0.00426
03-交响乐36 C大调“林兹”，K 425：小步舞曲：三重唱.wav
谐波能量
时间（s）
0.06
0.05
0.04
0.03
0.02
0.01
0
20
22
24
26
28
30
32
34
36
38
40

图 15

西班牙调	即兴精酿 唱片2	持续时间17.32
油	半个尼尔森 CD 9	持续时间5.52

再见黑鸟	再见黑鸟 CD 8	持续时间7.55
萨里郡的边缘上	四 CD 7	持续时间9.06

那又怎样	某种蓝调	持续时间9.22
艾哈迈德的蓝调	半个尼尔森 CD 9	持续时间7.27

当我坠入爱河	半个尼尔森 CD 9	持续时间4.25
我可笑的情人节	我可笑的情人节 CD 10	持续时间5.59

弗拉门戈素描	某种蓝调	持续时间9.26
绿中蓝	某种蓝调	持续时间5.73

图 16

分类	上升	平稳	下降
5.		•	•
4.	•	•	•
3.	•	•	•
2.	•	•	•
1.	•	•	

图 17

种类	上升	平稳	下降
5.		A大调第七交响曲，作品92 有活力的快板	A大调第七交响曲，作品92 急板不太快的急板
4.	C小调第七交响曲，作品67 有活力的快板		F大调第八交响曲，作品93 小步舞曲
3.	D小调第九交响曲‘欢乐颂’，作品125 庄严的快板	F大调第六交响曲 ‘帕斯朵拉’，作品68 初到乡村时的愉快感受	A大调第七交响曲，作品92 小快板
2.	降E调第三交响曲 ‘英雄交响曲’，作品55 有活力的快板		F大调第六交响曲 ‘帕斯朵拉’，作品68 牧羊人的歌声：暴风雨过后的愉快和感恩的情绪
1.	降E调第三交响曲 ‘英雄交响曲’，作品55 葬礼进行曲：很慢的柔板	F大调第六交响曲 ‘帕斯朵拉’，作品68 场景在小溪	

图 18