US006862558B2

(12) **United States Patent**
Huang

(10) **Patent No.:** **US 6,862,558 B2**
(45) **Date of Patent:** **Mar. 1, 2005**

(54) **EMPIRICAL MODE DECOMPOSITION FOR ANALYZING ACOUSTICAL SIGNALS**

(75) Inventor: **Norden E. Huang**, Bethesda, MD (US)

(73) Assignee: **The United States of America as represented by the Administrator of the National Aeronautics and Space Administration**, Washington, DC (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 141 days.

(21) Appl. No.: **10/073,957**

(22) Filed: **Feb. 13, 2002**

(65) **Prior Publication Data**

US 2003/0033094 A1 Feb. 13, 2003

Related U.S. Application Data

(63) Continuation-in-part of application No. 10/011,206, filed on Dec. 10, 2001.
(60) Provisional application No. 60/268,422, filed on Feb. 14, 2001.

(51) **Int. Cl.**[7] **H04B 15/00**; G06F 19/00
(52) **U.S. Cl.** **702/194**; 702/179; 708/300; 704/226
(58) **Field of Search** 702/66, 39, 106, 702/124, 3–17, 20–23, 70, 75–78, 126, 179–181, 189–191, 193–199; 73/579, 594; 708/300, 303, 304, 317, 320; 704/200, 201, 205, 221, 226, 231, 233

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,747,056 A * 7/1973 Treybig et al. 367/62

5,983,162 A * 11/1999 Huang 702/4
6,192,758 B1 2/2001 Huang
6,381,559 B1 4/2002 Huang

OTHER PUBLICATIONS

Timashev, S. A., and Shalin, M. G., 2000: Application of the Numerical Hilbert Translform to machine vibrodiagnostics. Proc. of COMADEM 2000, 175–185.

* cited by examiner

Primary Examiner—Marc S. Hoff
Assistant Examiner—Manuel L. Barbee
(74) *Attorney, Agent, or Firm*—Keith L. Dixon

(57) **ABSTRACT**

The present invention discloses a computer implemented signal analysis method through the Hilbert-Huang Transformation (HHT) for analyzing acoustical signals, which are assumed to be nonlinear and nonstationary. The Empirical Decomposition Method (EMD) and the Hilbert Spectral Analysis (HSA) are used to obtain the HHT. Essentially, the acoustical signal will be decomposed into the Intrinsic Mode Function Components (IMFs). Once the invention decomposes the acoustic signal into its constituting components, all operations such as analyzing, identifying, and removing unwanted signals can be performed on these components. Upon transforming the IMFs into Hilbert spectrum, the acoustical signal may be compared with other acoustical signals.

19 Claims, 121 Drawing Sheets

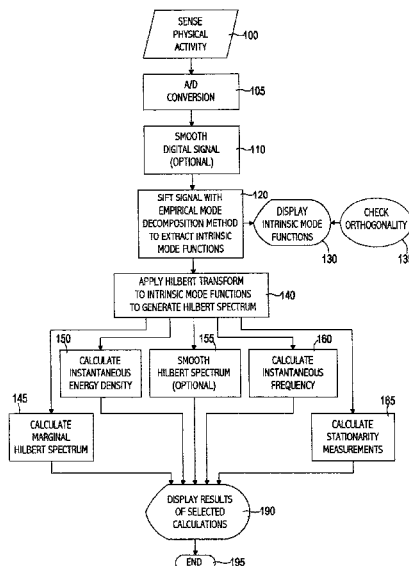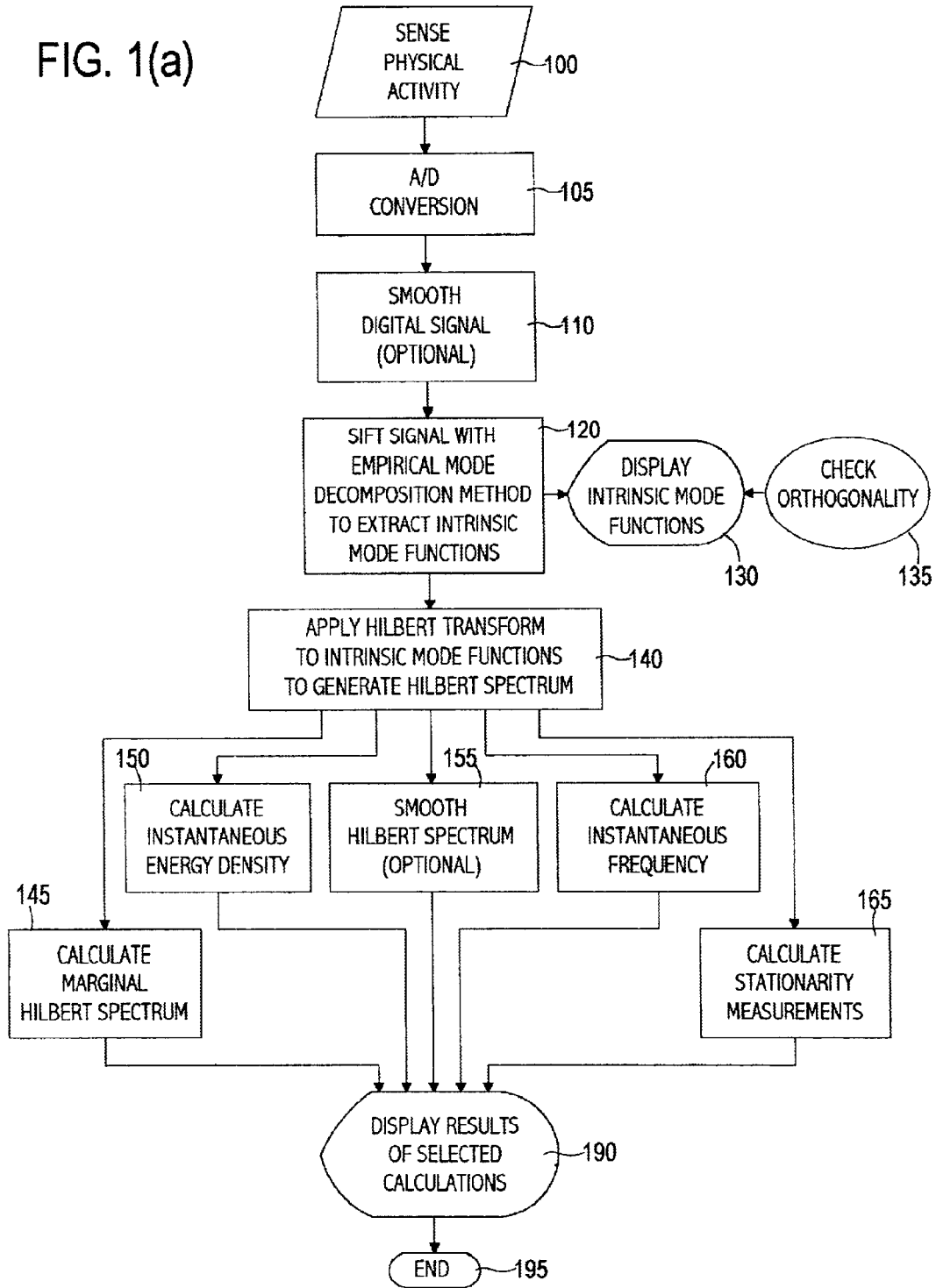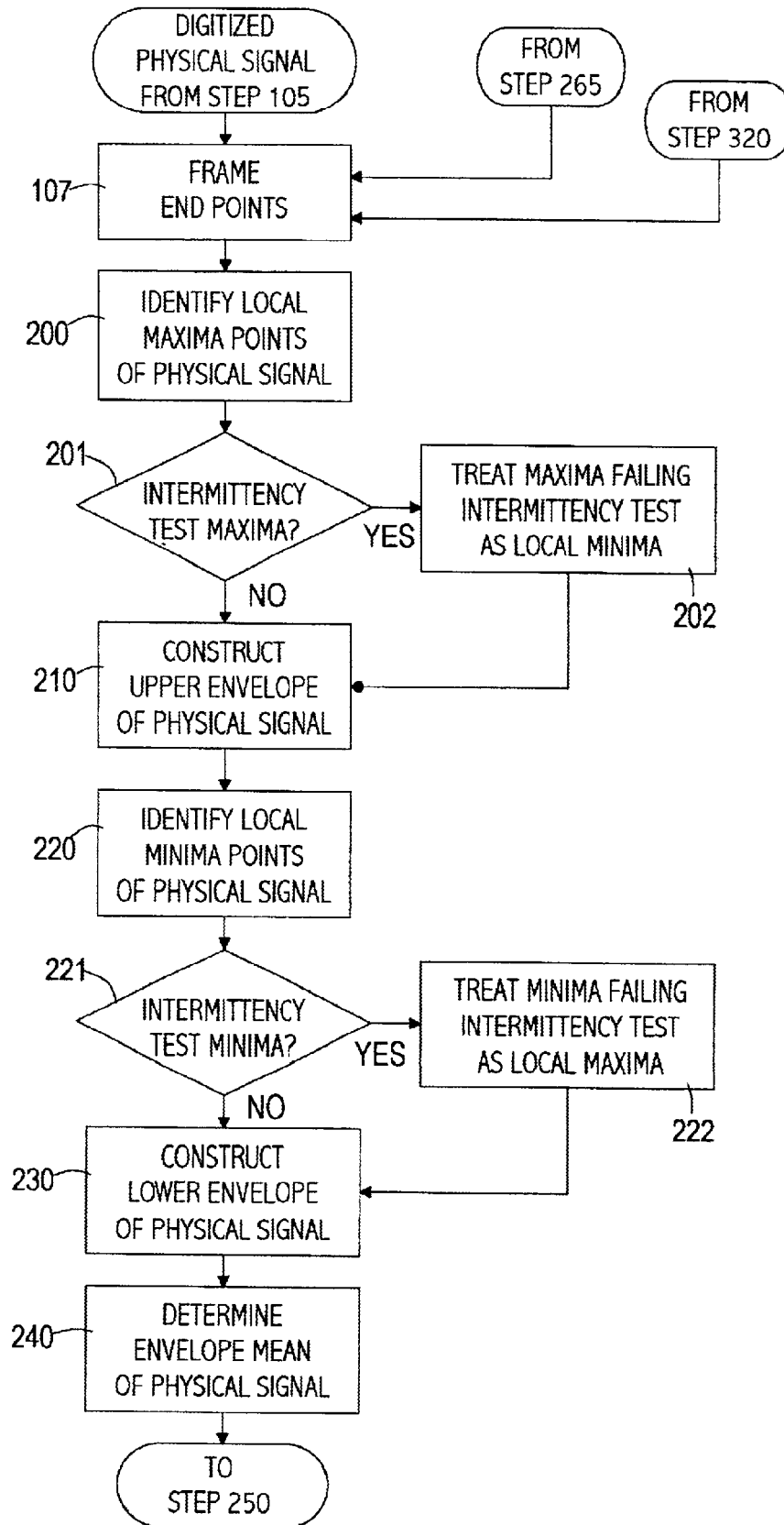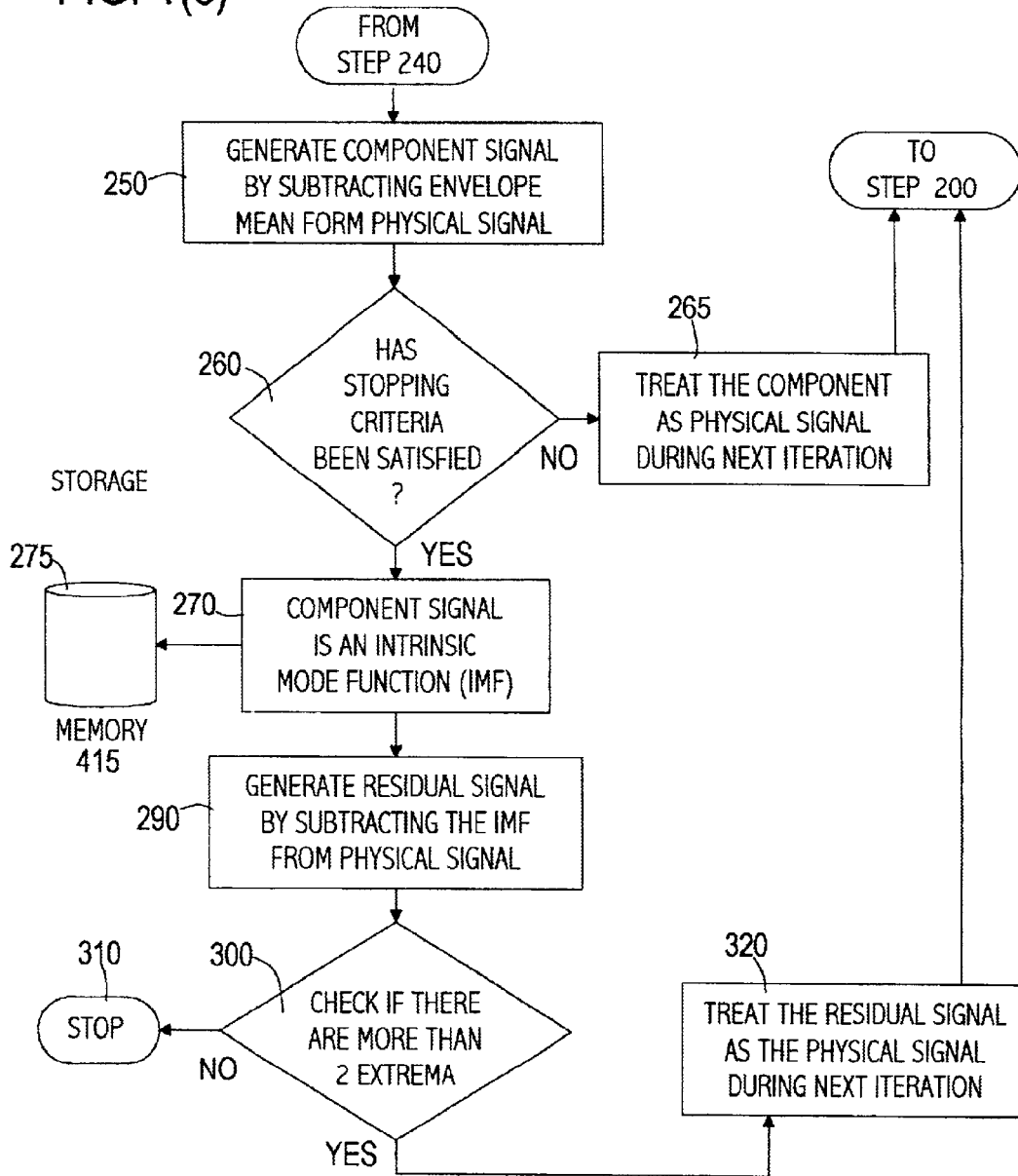**(27 of 121 Drawing Sheet(s) Filed in Color)**

FIG. 1(a)

SENSE PHYSICAL ACTIVITY —100

↓

A/D CONVERSION —105

↓

SMOOTH DIGITAL SIGNAL (OPTIONAL) —110

↓

SIFT SIGNAL WITH EMPIRICAL MODE DECOMPOSITION METHOD TO EXTRACT INTRINSIC MODE FUNCTIONS —120 → DISPLAY INTRINSIC MODE FUNCTIONS 130 ← CHECK ORTHOGONALITY 135

↓

APPLY HILBERT TRANSFORM TO INTRINSIC MODE FUNCTIONS TO GENERATE HILBERT SPECTRUM —140

150 — CALCULATE INSTANTANEOUS ENERGY DENSITY

155 — SMOOTH HILBERT SPECTRUM (OPTIONAL)

160 — CALCULATE INSTANTANEOUS FREQUENCY

145 — CALCULATE MARGINAL HILBERT SPECTRUM

165 — CALCULATE STATIONARITY MEASUREMENTS

↓

DISPLAY RESULTS OF SELECTED CALCULATIONS —190

↓

END —195

FIG. 1(b)

DIGITIZED
PHYSICAL SIGNAL
FROM STEP 105

FROM
STEP 265

FROM
STEP 320

107 — FRAME
END POINTS

200 — IDENTIFY LOCAL
MAXIMA POINTS
OF PHYSICAL SIGNAL

201 — INTERMITTENCY
TEST MAXIMA? → YES → TREAT MAXIMA FAILING
INTERMITTENCY TEST
AS LOCAL MINIMA

202

NO

210 — CONSTRUCT
UPPER ENVELOPE
OF PHYSICAL SIGNAL

220 — IDENTIFY LOCAL
MINIMA POINTS
OF PHYSICAL SIGNAL

221 — INTERMITTENCY
TEST MINIMA? → YES → TREAT MINIMA FAILING
INTERMITTENCY TEST
AS LOCAL MAXIMA

222

NO

230 — CONSTRUCT
LOWER ENVELOPE
OF PHYSICAL SIGNAL

240 — DETERMINE
ENVELOPE MEAN
OF PHYSICAL SIGNAL

TO
STEP 250

FIG. 1(c)

FROM
STEP 240

250 — GENERATE COMPONENT SIGNAL
BY SUBTRACTING ENVELOPE
MEAN FORM PHYSICAL SIGNAL

TO
STEP 200

260 — HAS
STOPPING
CRITERIA
BEEN SATISFIED
?

NO

265

TREAT THE COMPONENT
AS PHYSICAL SIGNAL
DURING NEXT ITERATION

STORAGE

YES

275

270 — COMPONENT SIGNAL
IS AN INTRINSIC
MODE FUNCTION (IMF)

MEMORY
415

290 — GENERATE RESIDUAL SIGNAL
BY SUBTRACTING THE IMF
FROM PHYSICAL SIGNAL

310

300 — CHECK IF THERE
ARE MORE THAN
2 EXTREMA

STOP

NO

320

TREAT THE RESIDUAL SIGNAL
AS THE PHYSICAL SIGNAL
DURING NEXT ITERATION

YES

```
        ┌─────────────────────────┐
       /        SENSE             /──100
      /  PHYSICAL PHENOMENON     /
     /          OR              /
    /   INPUT PHYSICAL SIGNAL  /
    └─────────────────────────┘
                │
                ▼
       ┌──────────────────┐
       │  SMOOTH SIGNAL   │──105
       │   (OPTIONAL)     │
       └──────────────────┘
                │                                    ┌──────────┐
          120   │                                    │ DISPLAY  │──130
                ▼                                     └──────────┘
       ┌──────────────────────────┐              ┌──────────┐
       │     SIFT SIGNAL WITH     │─────────────▶│  STORE   │──132
  ────▶│ EMPIRICAL MODE DECOPOSITION             └──────────┘
       │     TO EXTRACT IMF'S     │              ┌──────────┐
       └──────────────────────────┘─────────────▶│ TRANSMIT │
 370          │                                   └──────────┘
  ┌──────────────┐                                     │
  │  CALCULATE   │                                    134
  │ OSCILLATORY  │
  │ ENERGY ABOUT │
  │  THE MEAN    │
  └──────────────┘
         ▲        │
         │        ▼
       ┌──────────────────────────┐
       │ CONSTRUCT FILTERED SIGNAL│──350
       │  FROM A SUBSET OF IMF'S  │
       └──────────────────────────┘
                │
                ▼
       ┌──────────────────┐
       │  FIT A CURVE TO  │──360
       │ THE FILTERED SIGNAL
       └──────────────────┘
```

FIG. 1(d)

FIG. 2

400

480 — DISC

430 — KEYBOARD

420 — MASS STORAGE DEVICE

435 — MOUSE

440 — INPUT DEVICE

COMPUTER — 410

MEMORY — 415

442 — SENSOR

450 — DISPLAY

444 — SENSOR

460 — PRINTER

OUTPUT DEVICE — 470

FIG. 3(a)

FIG. 3(b)

FIG. 3(c)

First Component Signal (h_1)

FIG. 3(d)

FIG. 3(e)

Third Component Signal (h_{12})

FIG. 3(f)

First IMF Component

FIG. 3(g)

Data with Intermittency

EMD-IMF without intermittency option

FIG. 3(h)

c1

Time

FIG. 3(i)

c2

Time

FIG. 3(j)

c3

Time

FIG. 3(k)

EMD-IMF without intermittency option

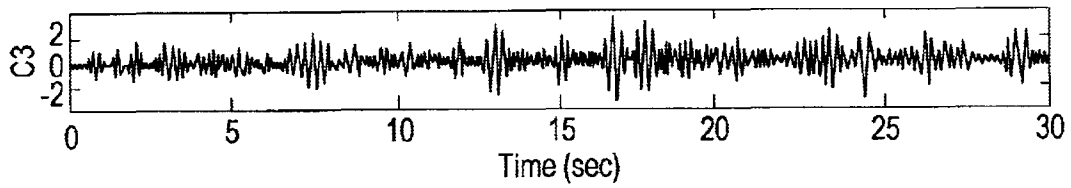FIG. 3(l)

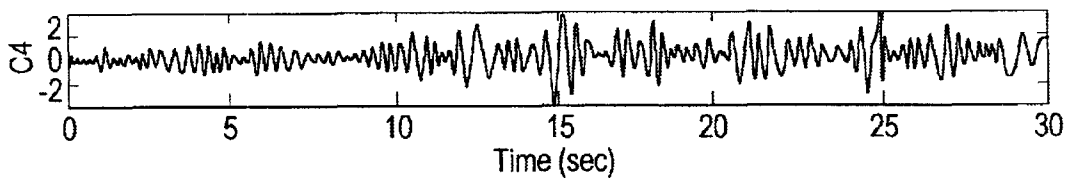FIG. 3(m)

FIG. 4(a)

Wind Data

FIG. 4(b)
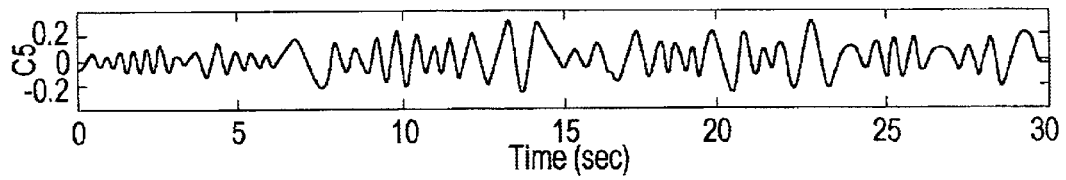

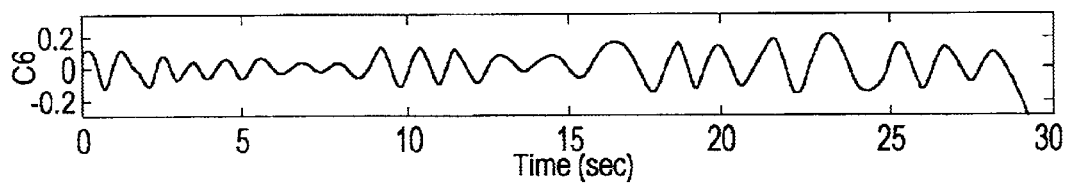
FIG. 4(c)

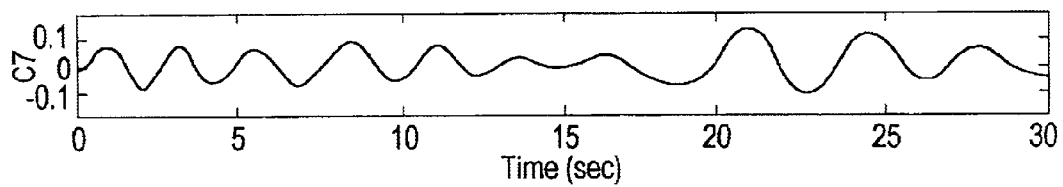

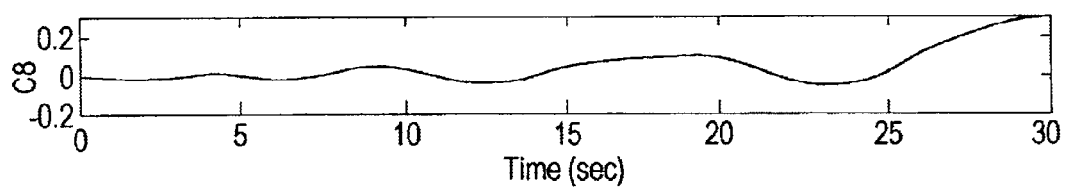FIG. 4(d)

FIG. 4(e)
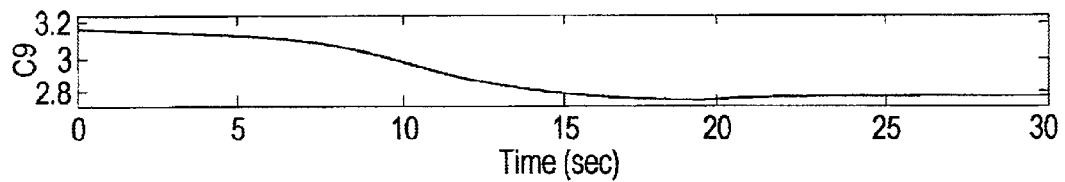

FIG. 4(f)

FIG. 4(g)

FIG. 4(h)

FIG. 4(i)

FIG. 4(j)

FIG. 4(k)

FIG. 5(a)

FIG. 5(b)

FIG. 5(c)

FIG. 5(d)

FIG. 5(e)

FIG. 5(f)

FIG. 5(g)

FIG. 5(h)

FIG. 5(i)

FIG. 5(j)

FIG. 6(a)
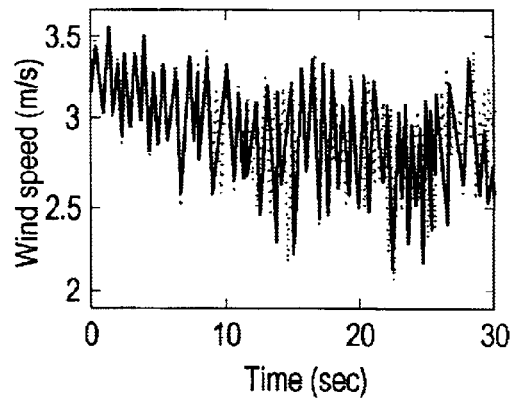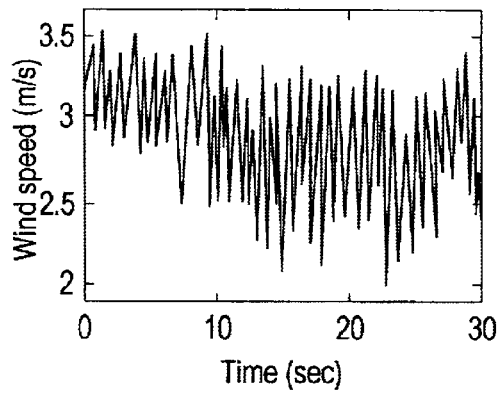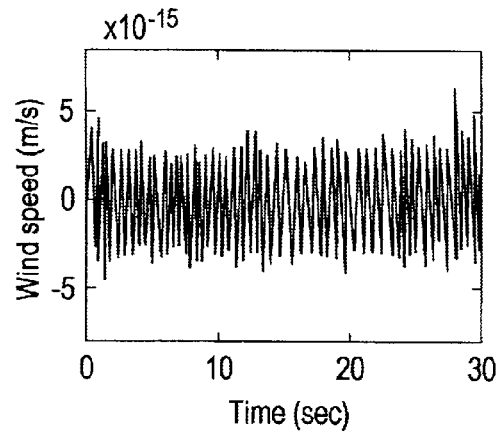
Hilbert Spectrum for Laboratory Wind Data
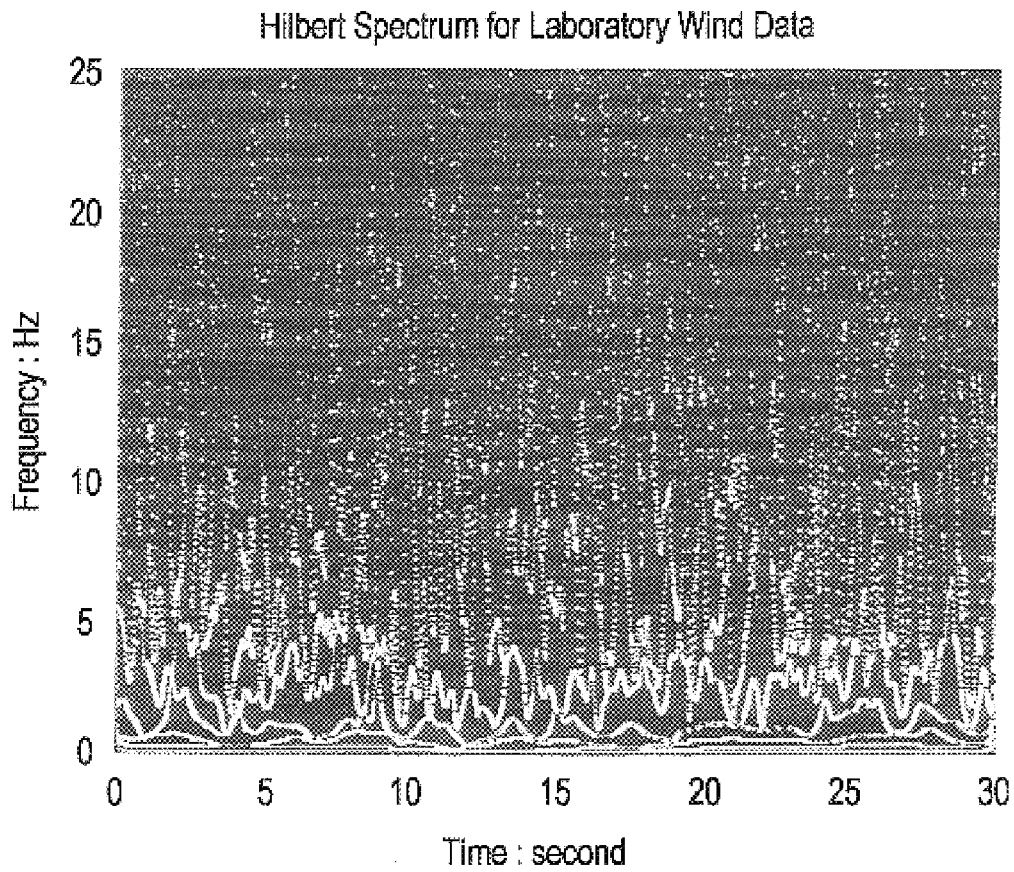
FIG. 6(b)

Wavelet Spectrum for Laboratory Wind Data

FIG. 6(c)

Hilbert Spectrum for Laboratory Wind Data : 15x15 Smoothed

Time : second

FIG. 7

FIG. 8(a)

FIG. 8(b)
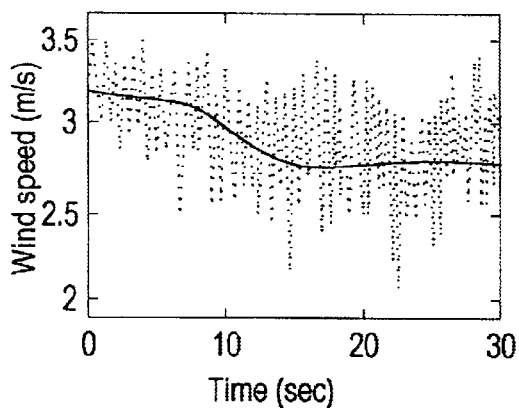

FIG. 8(c)

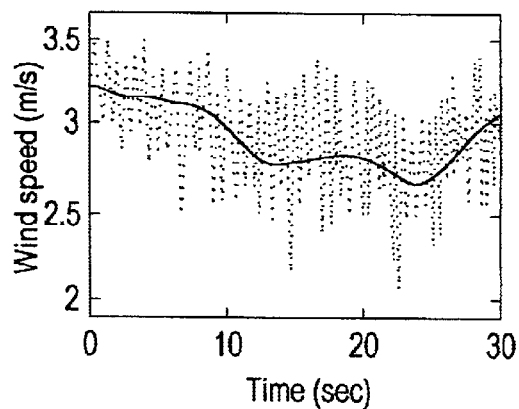PULMONARY ARTERIAL PRESSURE OF RAT (24 HR.)

FIG. 9(a)

ONE HOUR STRIP
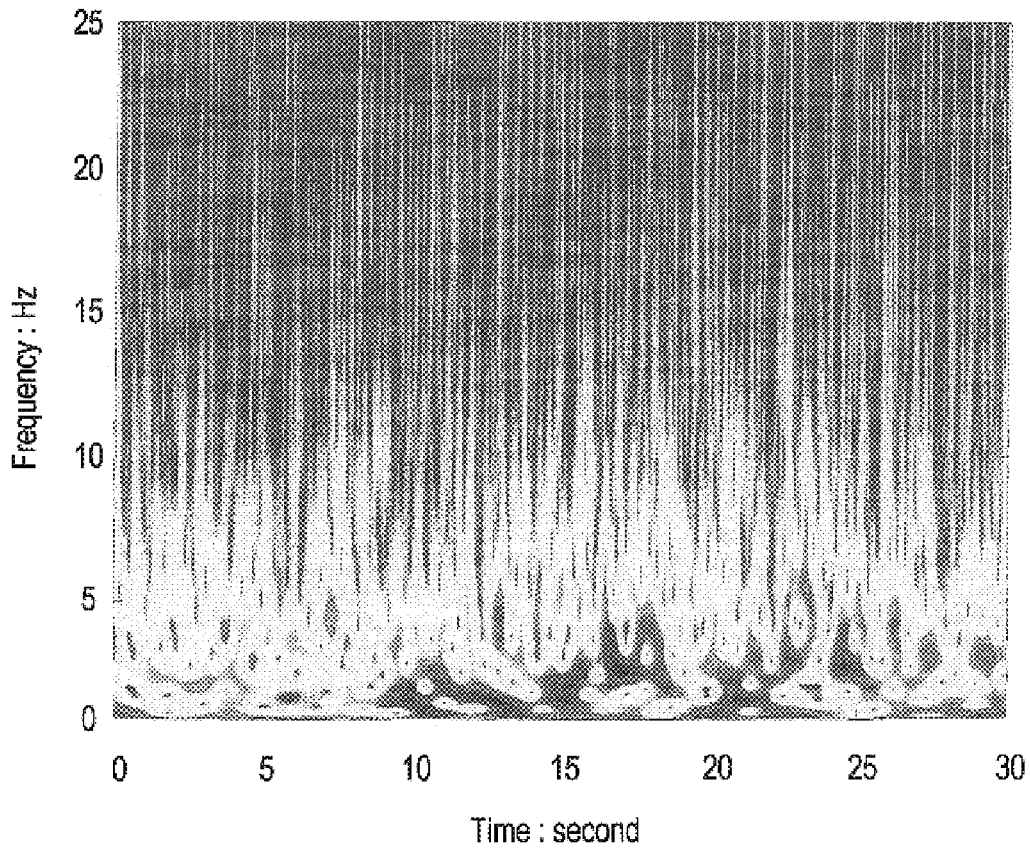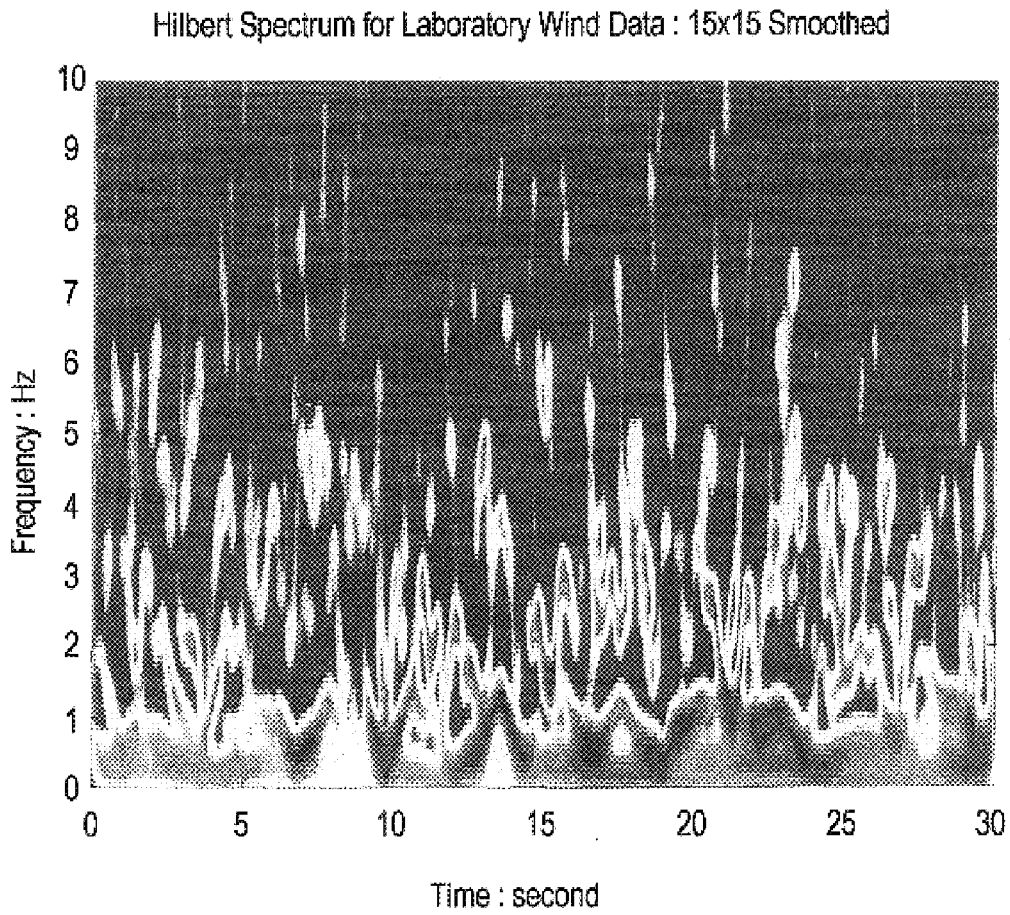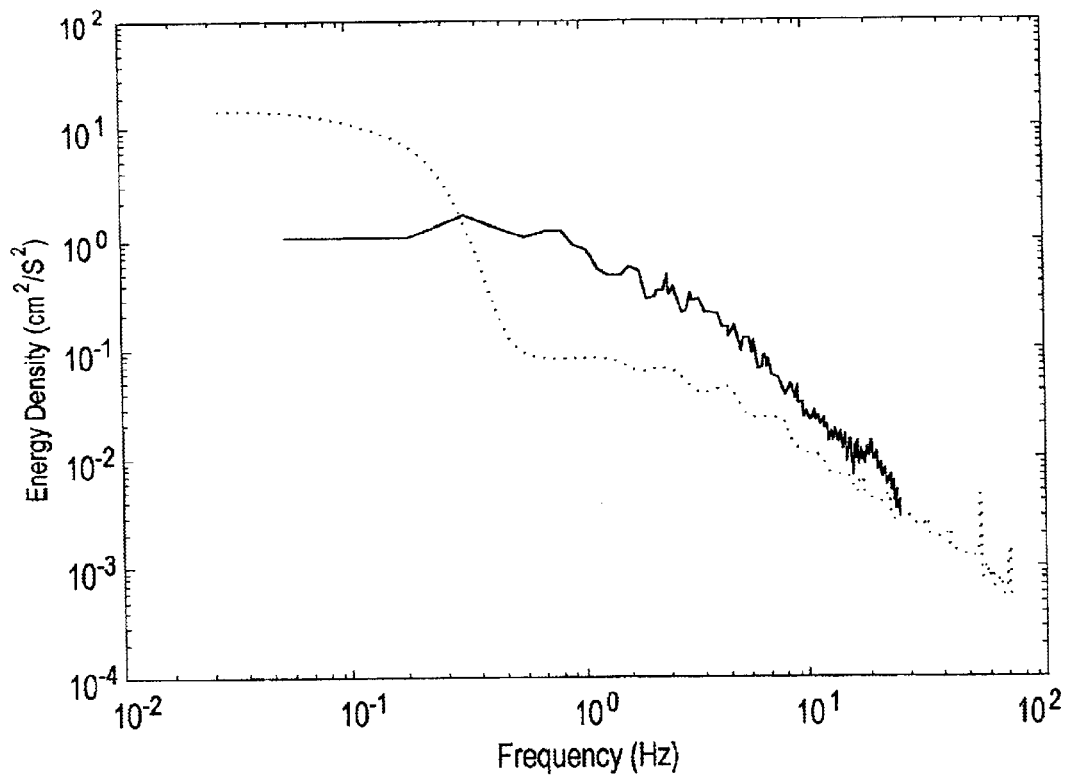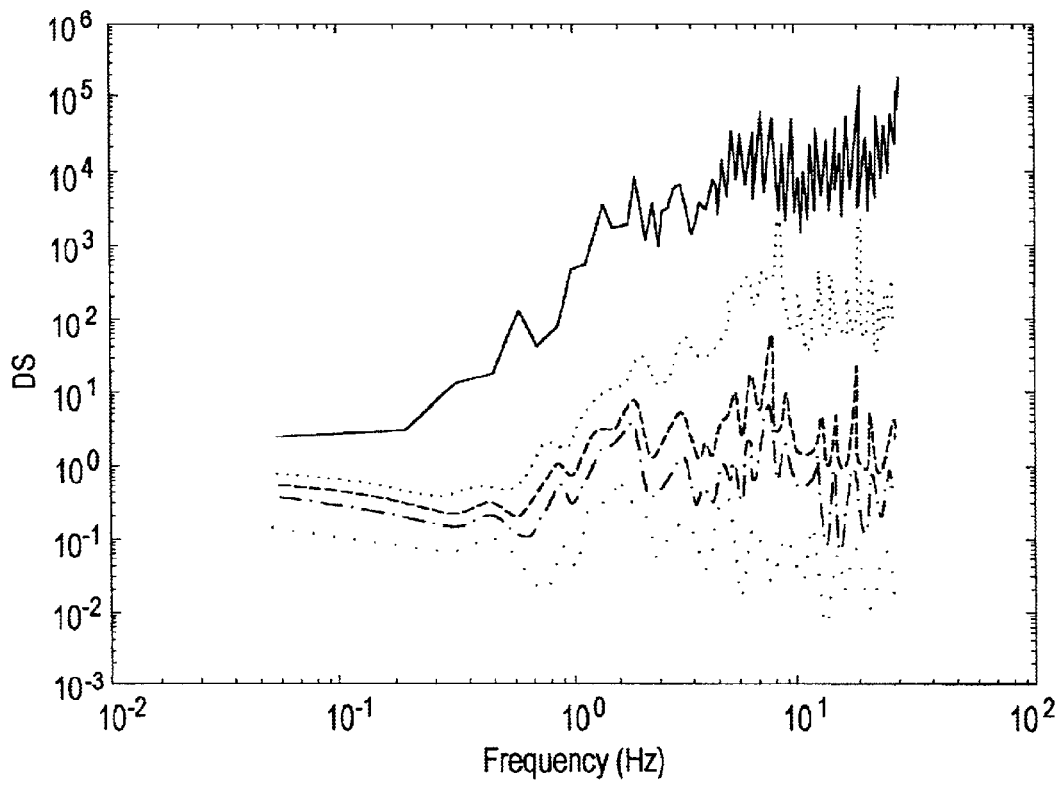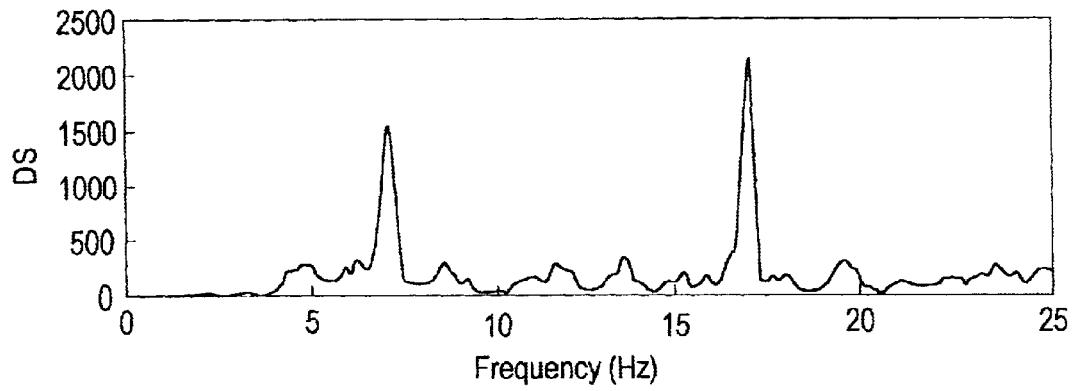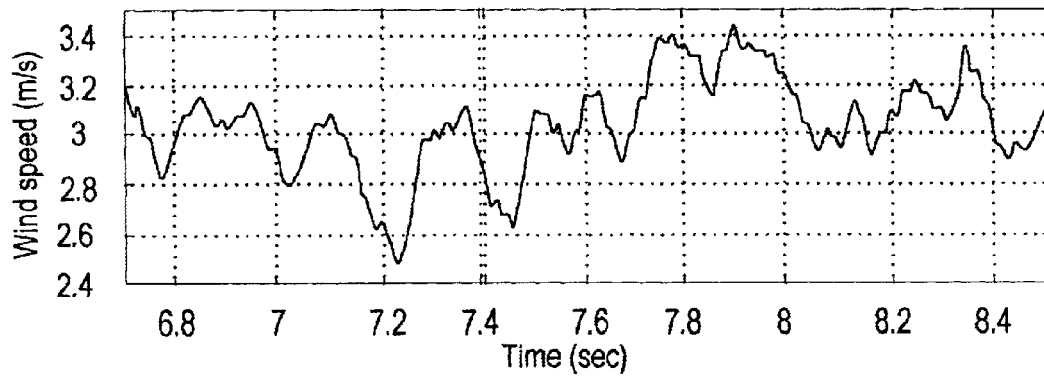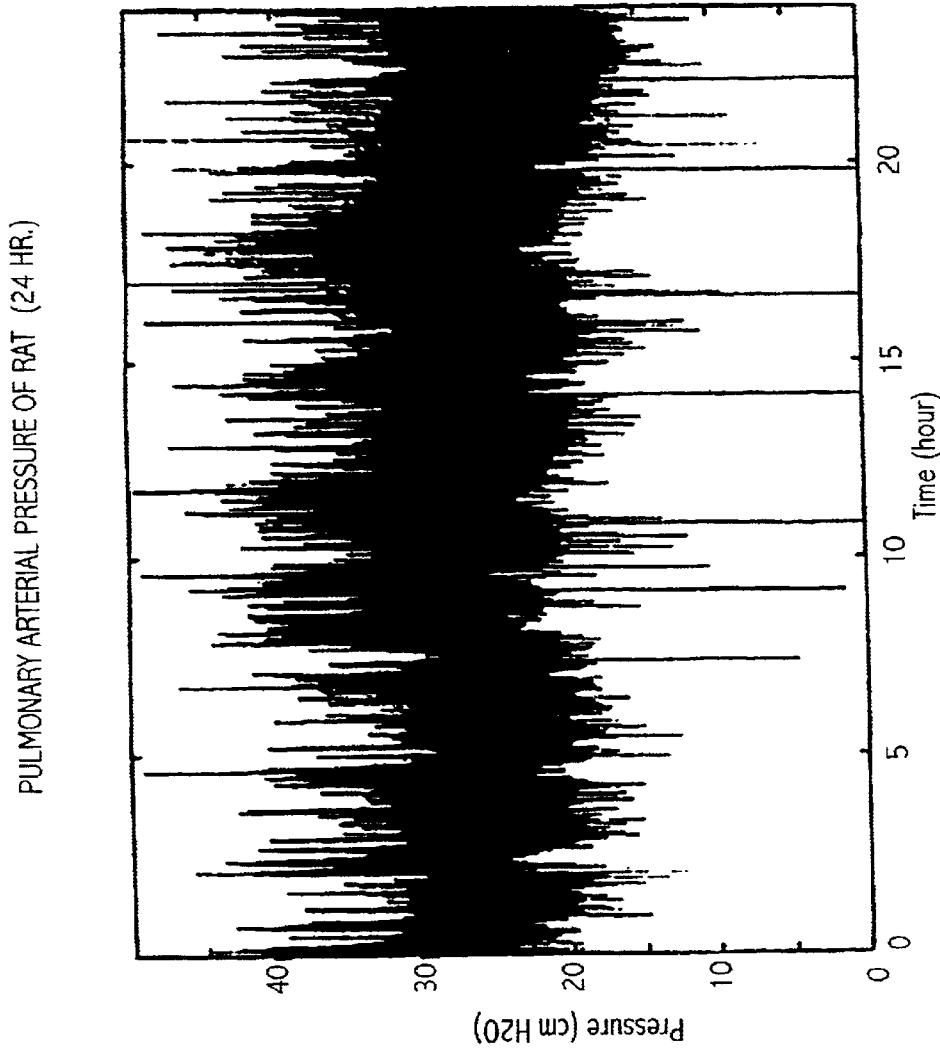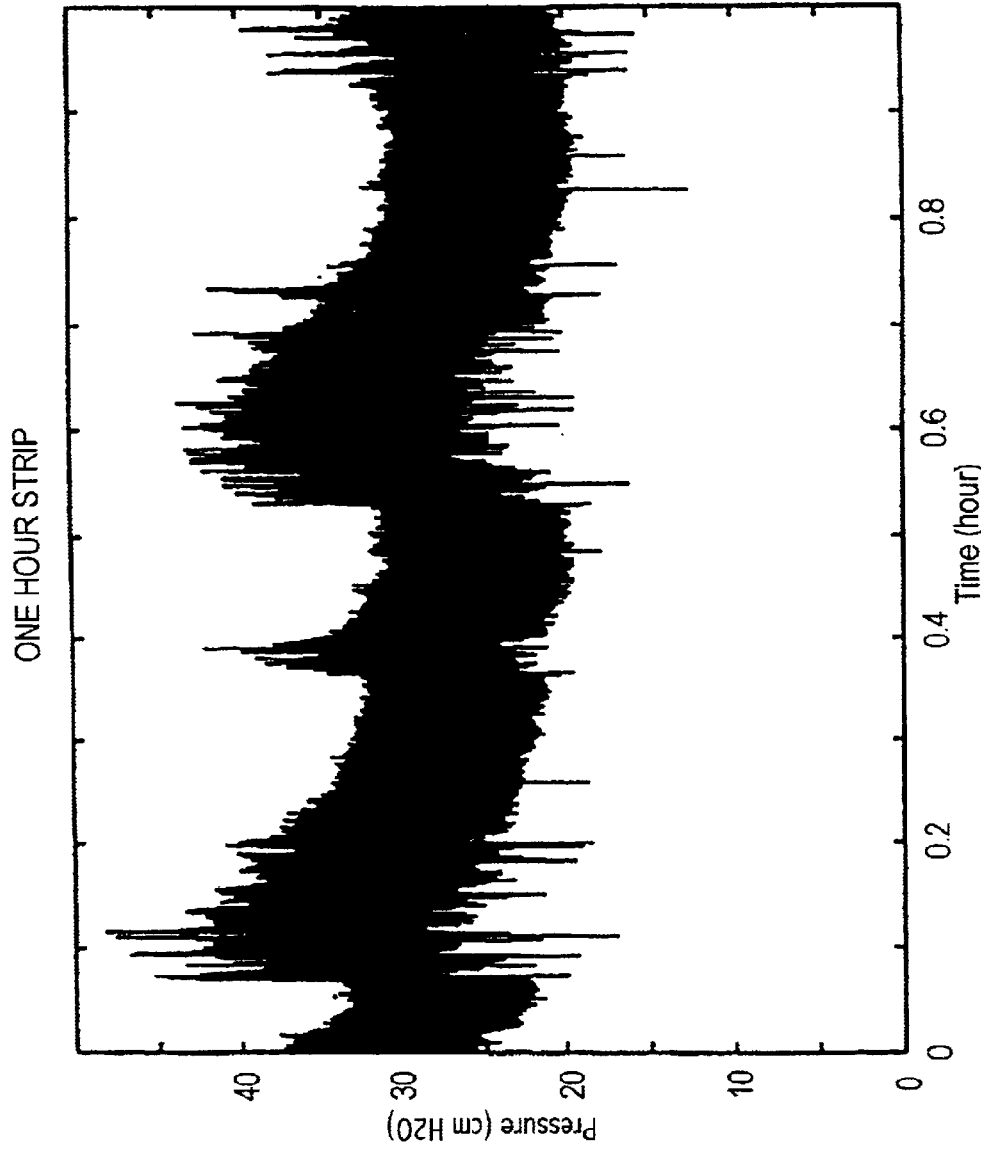
FIG. 9(b)

FIG. 9(c)

FIG. 9(d)

FIG. 9(e)

FIG. 9(f)

SYSTOLIC PRESSURES IN 1 HR.

FIG. 9(g)

DIASTOLIC PRESSURES IN 1 HR.
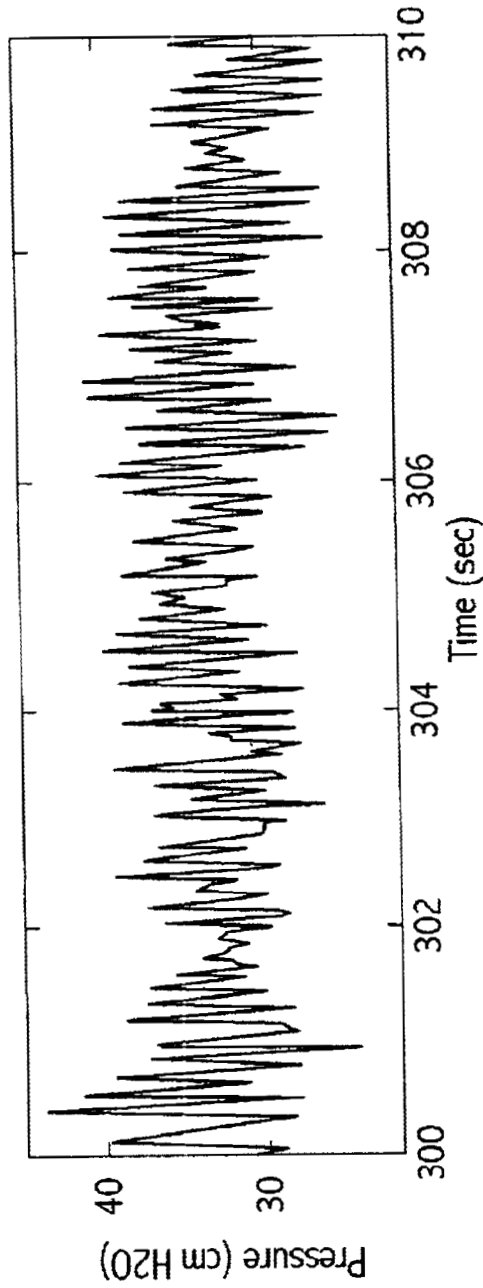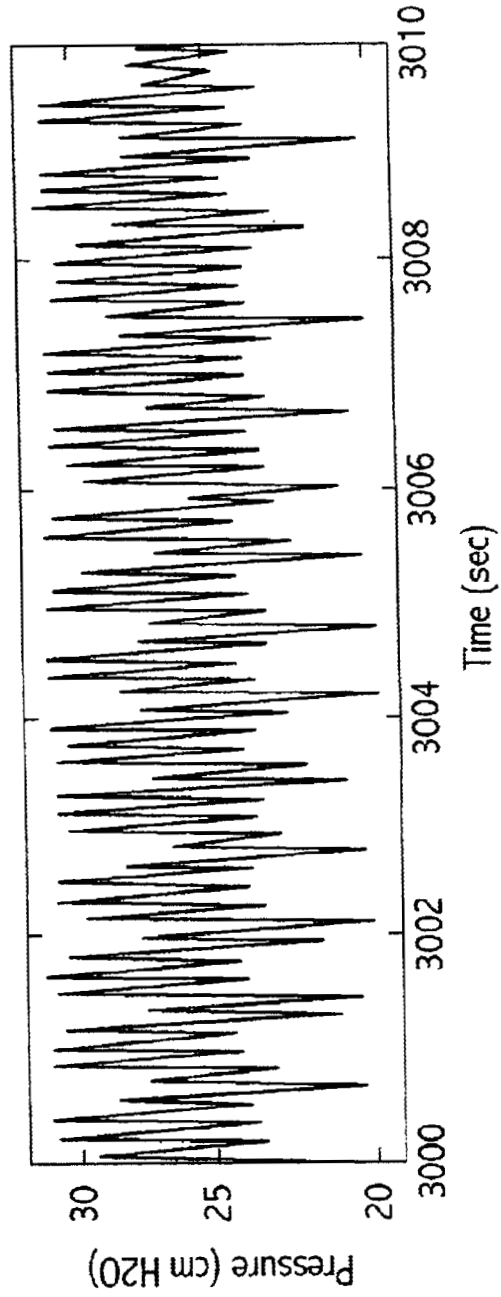
FIG. 10(a)
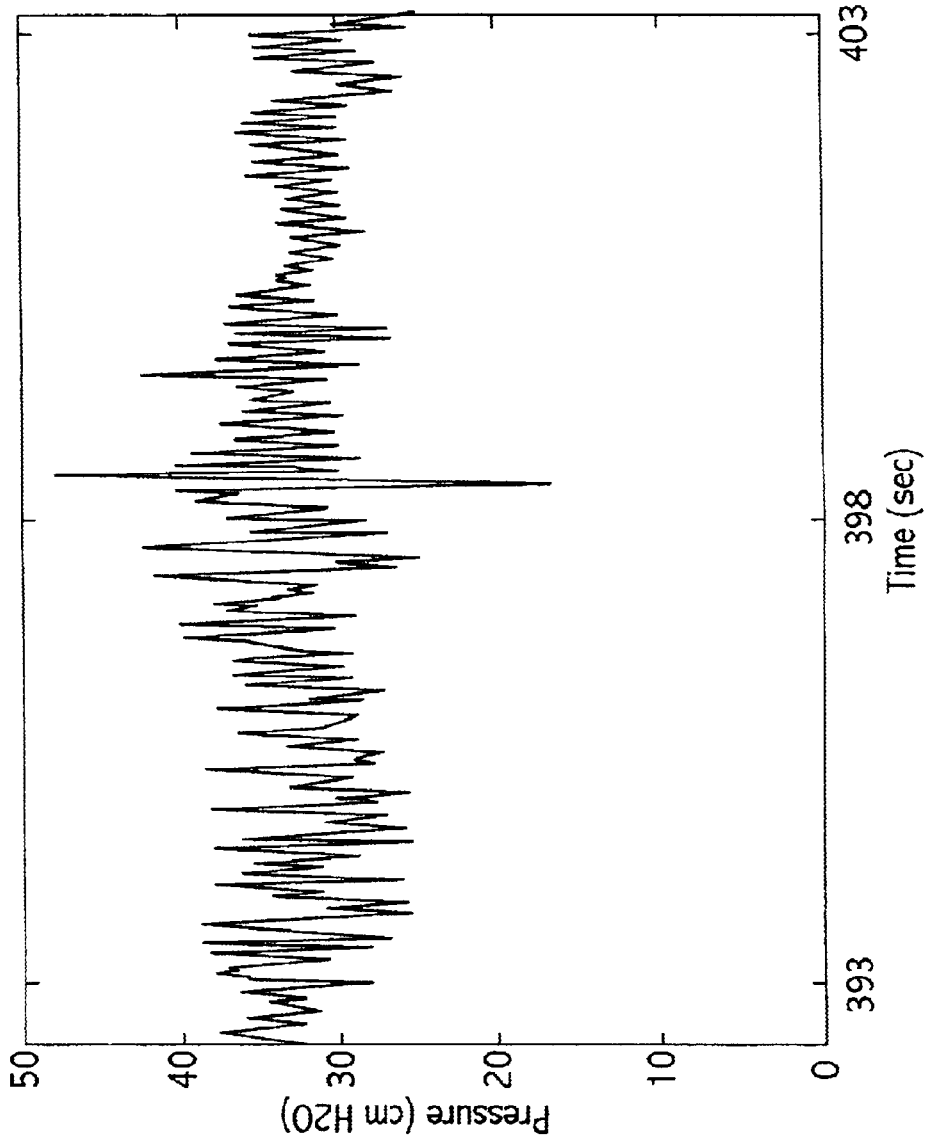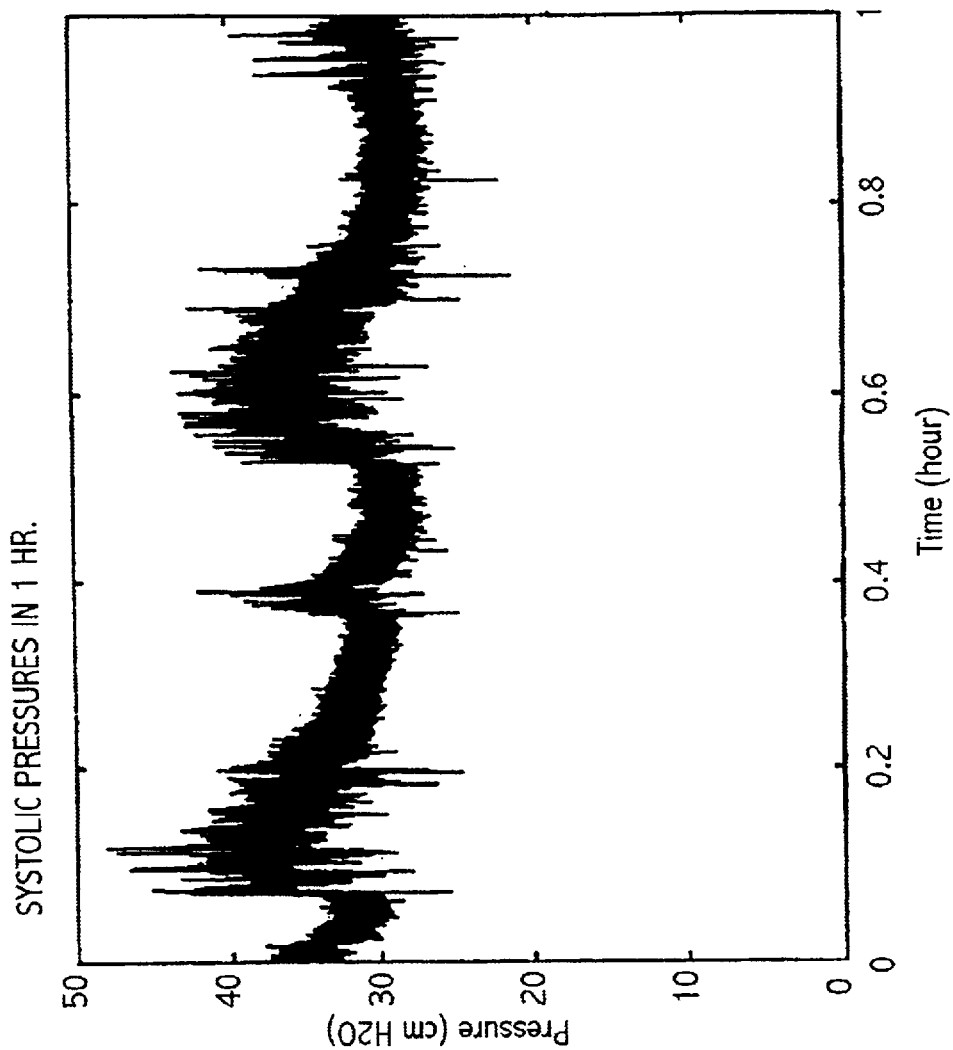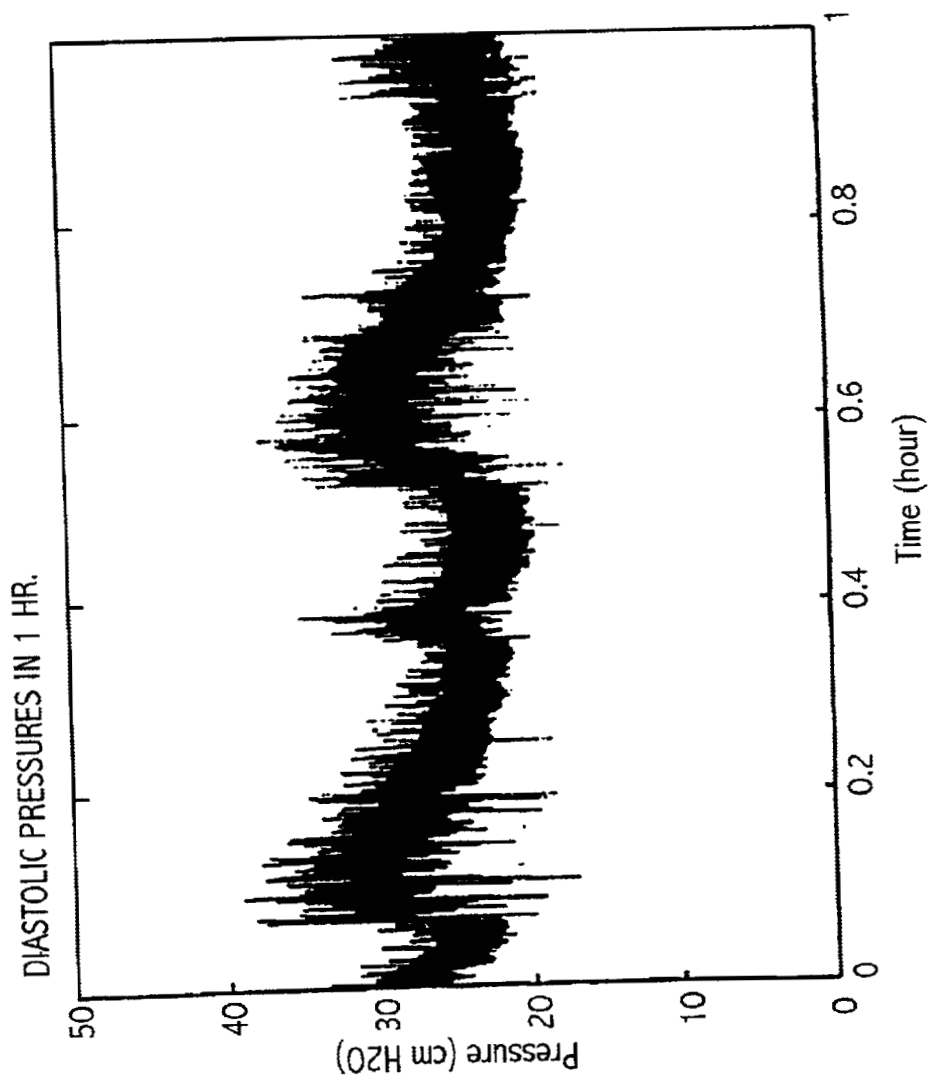
FOURIER SPECTRUM OF THE FIRST HR. DATA
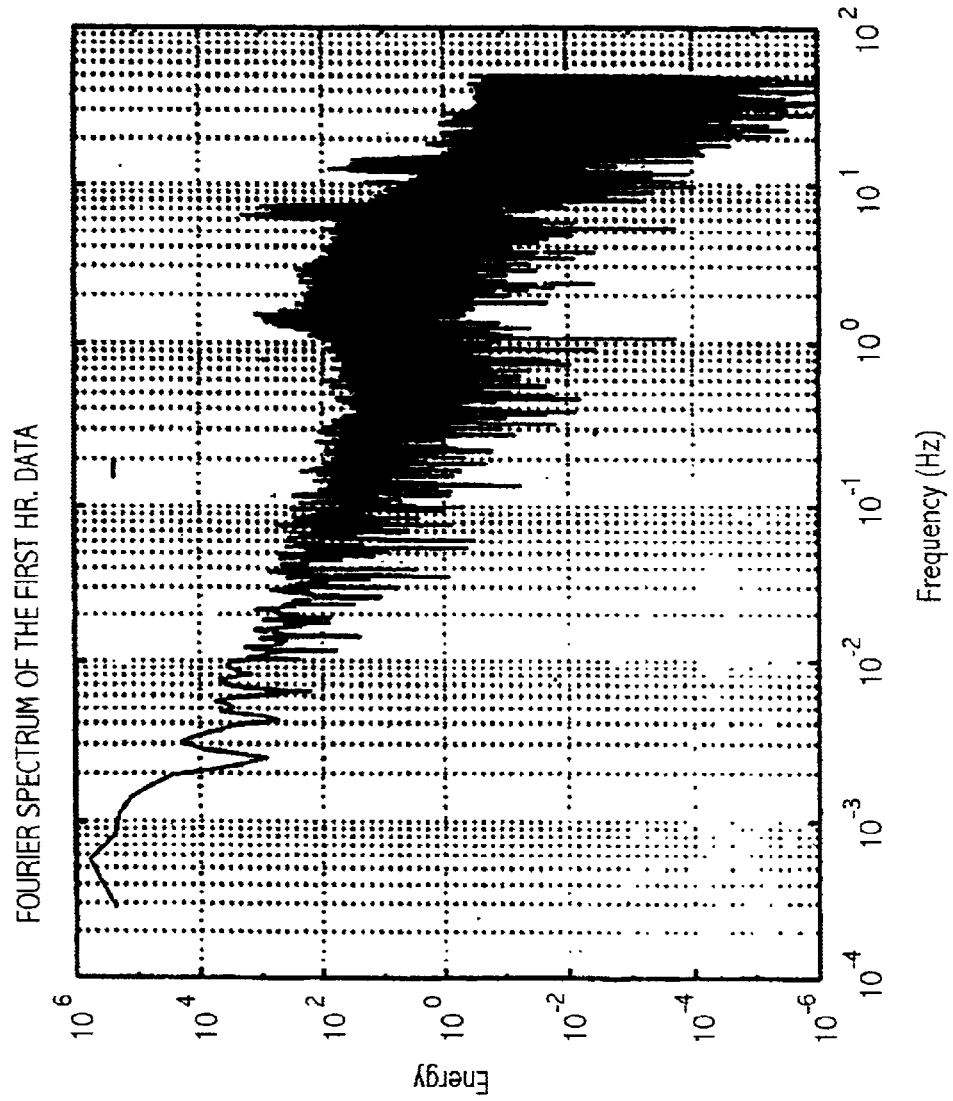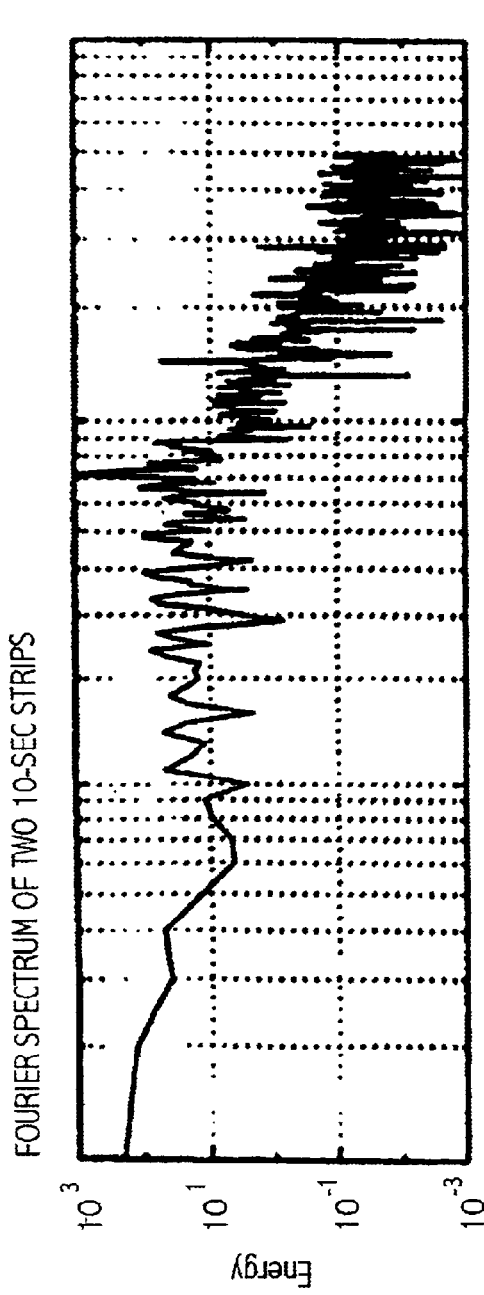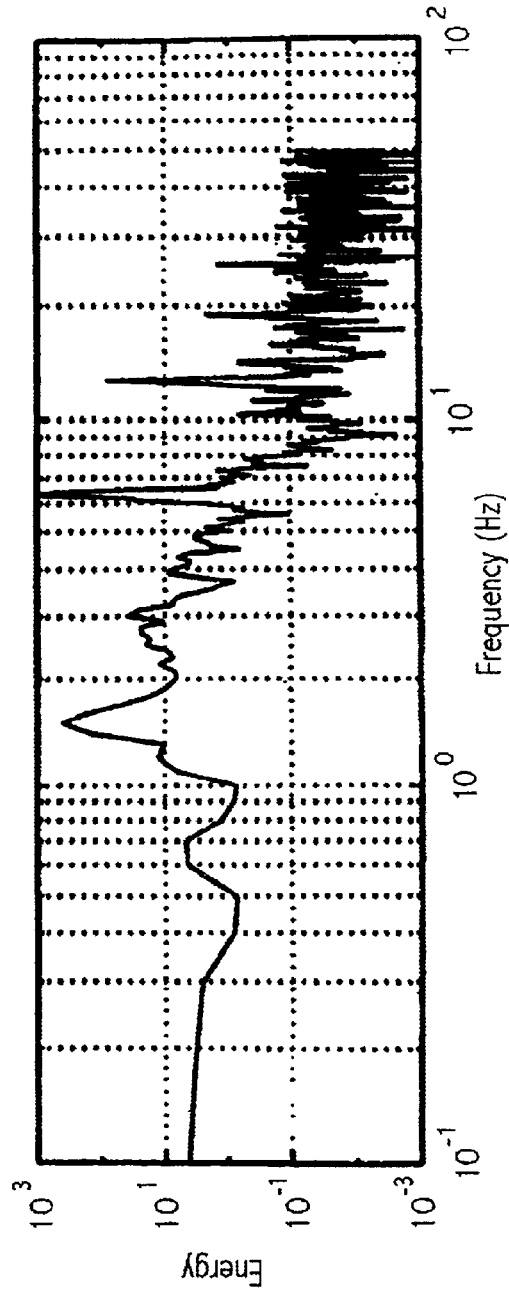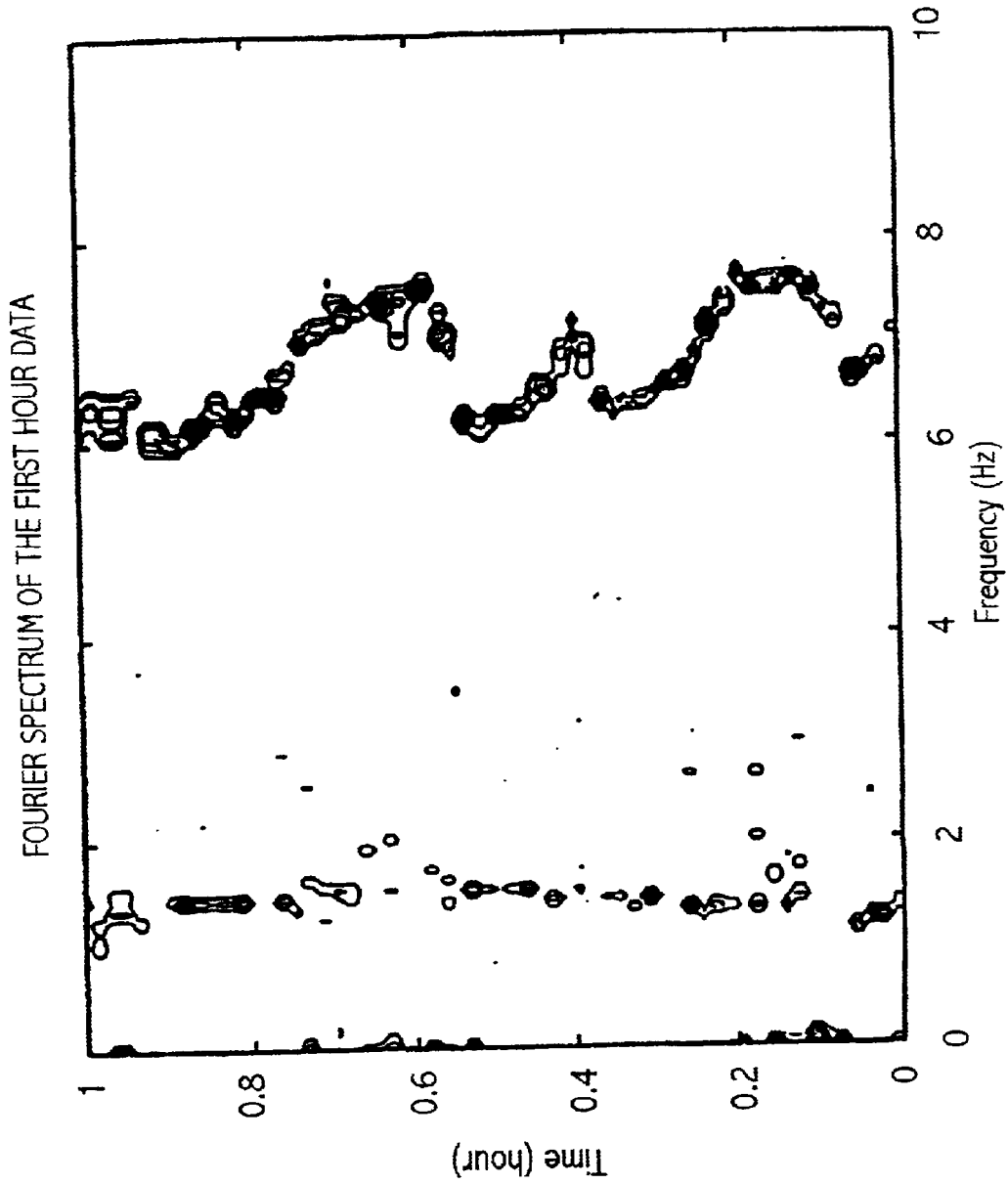
FIG. 10(b)

FIG. 10(c)

FIG. 10(d)

FIG. 10(e)

FIG. 10(f)

FOURIER AND MARGINAL HILBERT SPECTRA, FIRST 10-SEC.
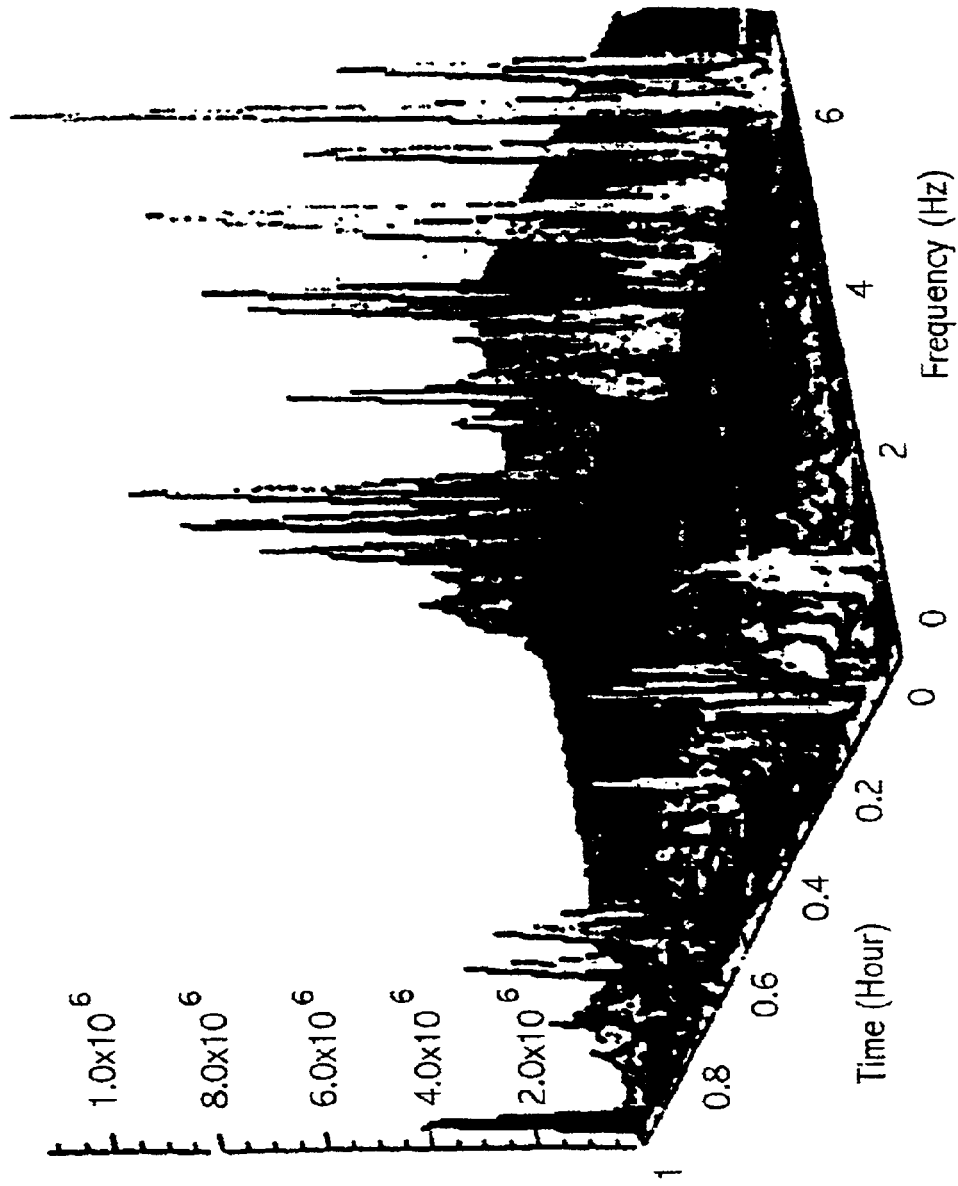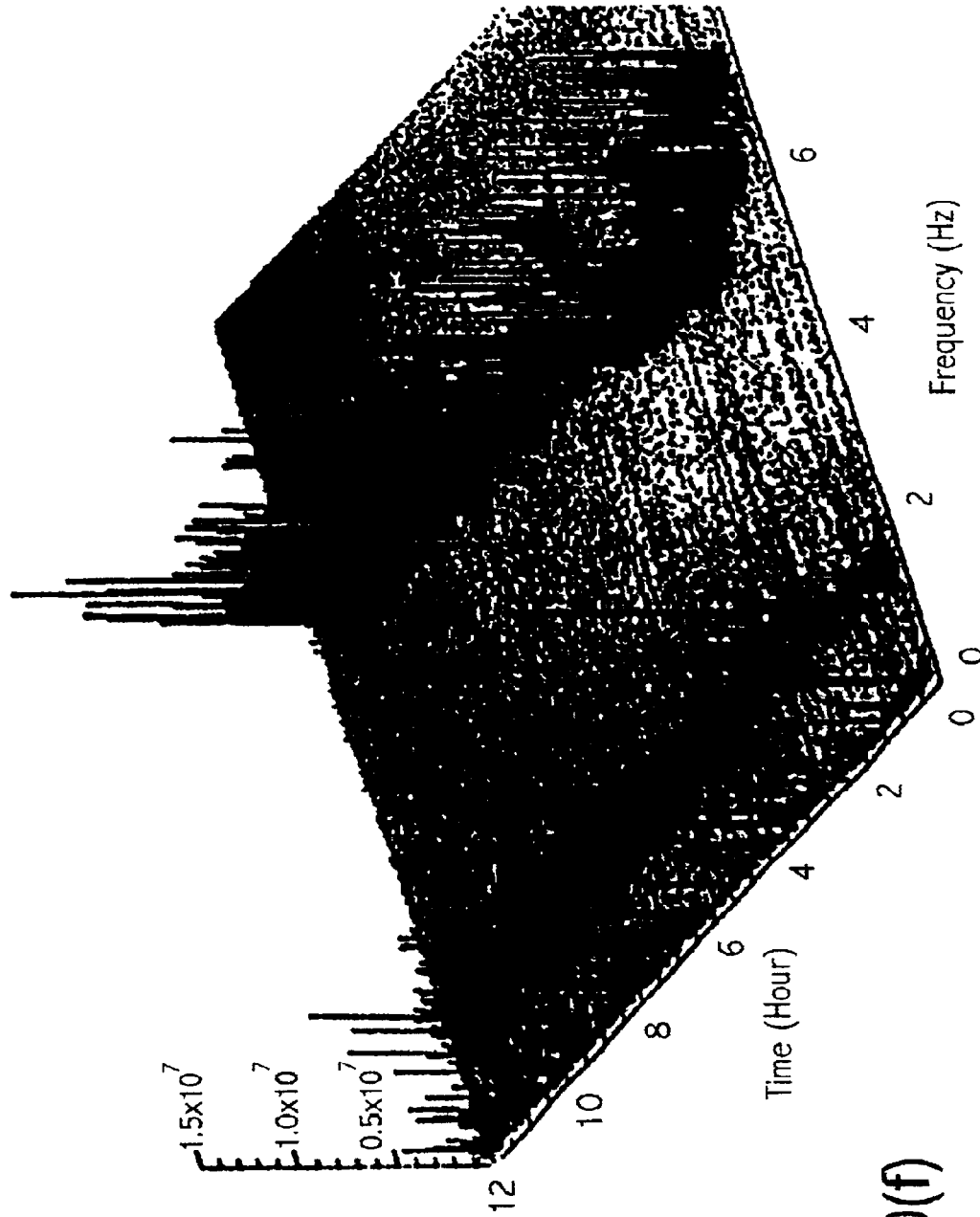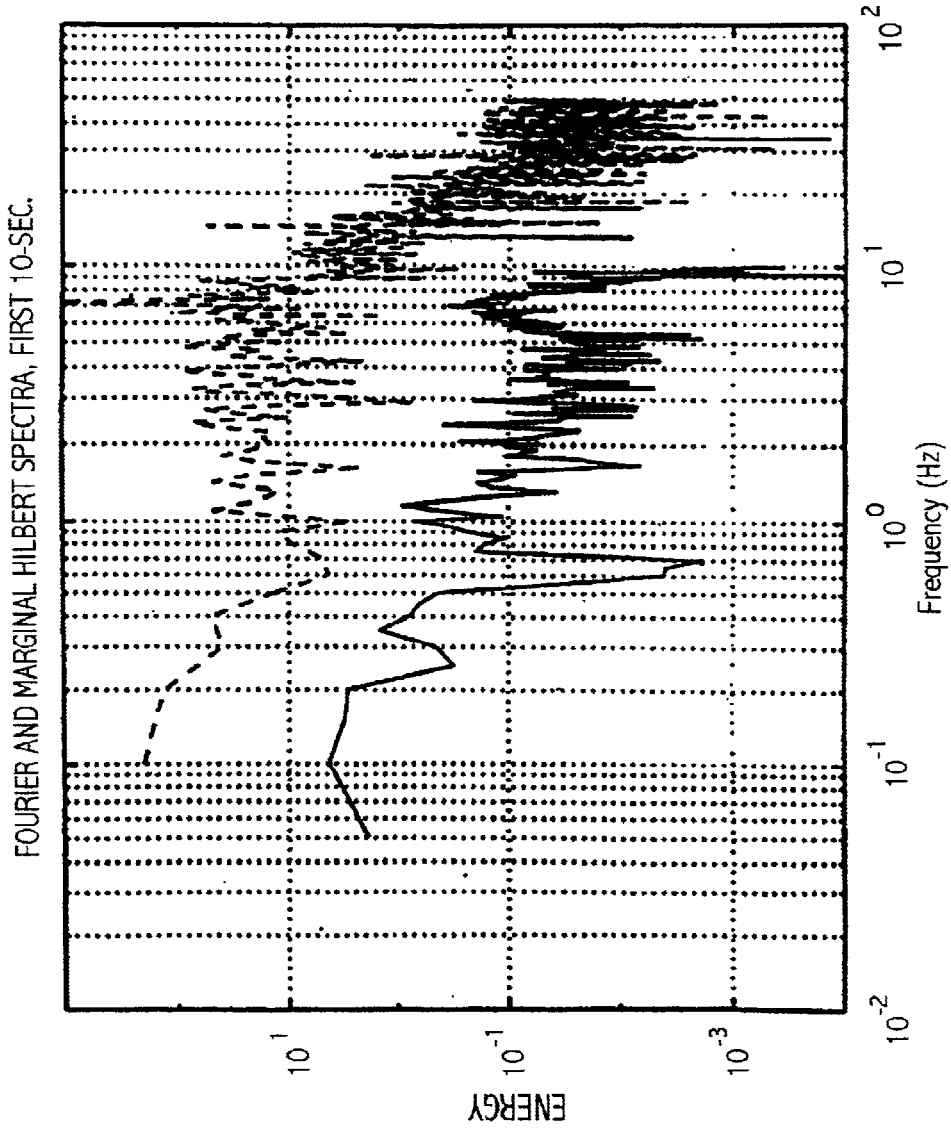
FIG. 10(g)

IMF-COMPONENTS OF THE FIRST 10-SEC. STRIP

FIG. 11(a)

FIG. 11(b)

FIG. 11(c)

FIG. 11(d)

FIG. 11(e)

FIG. 11(f)

FIG. 11(g)

FIG. 11(h)
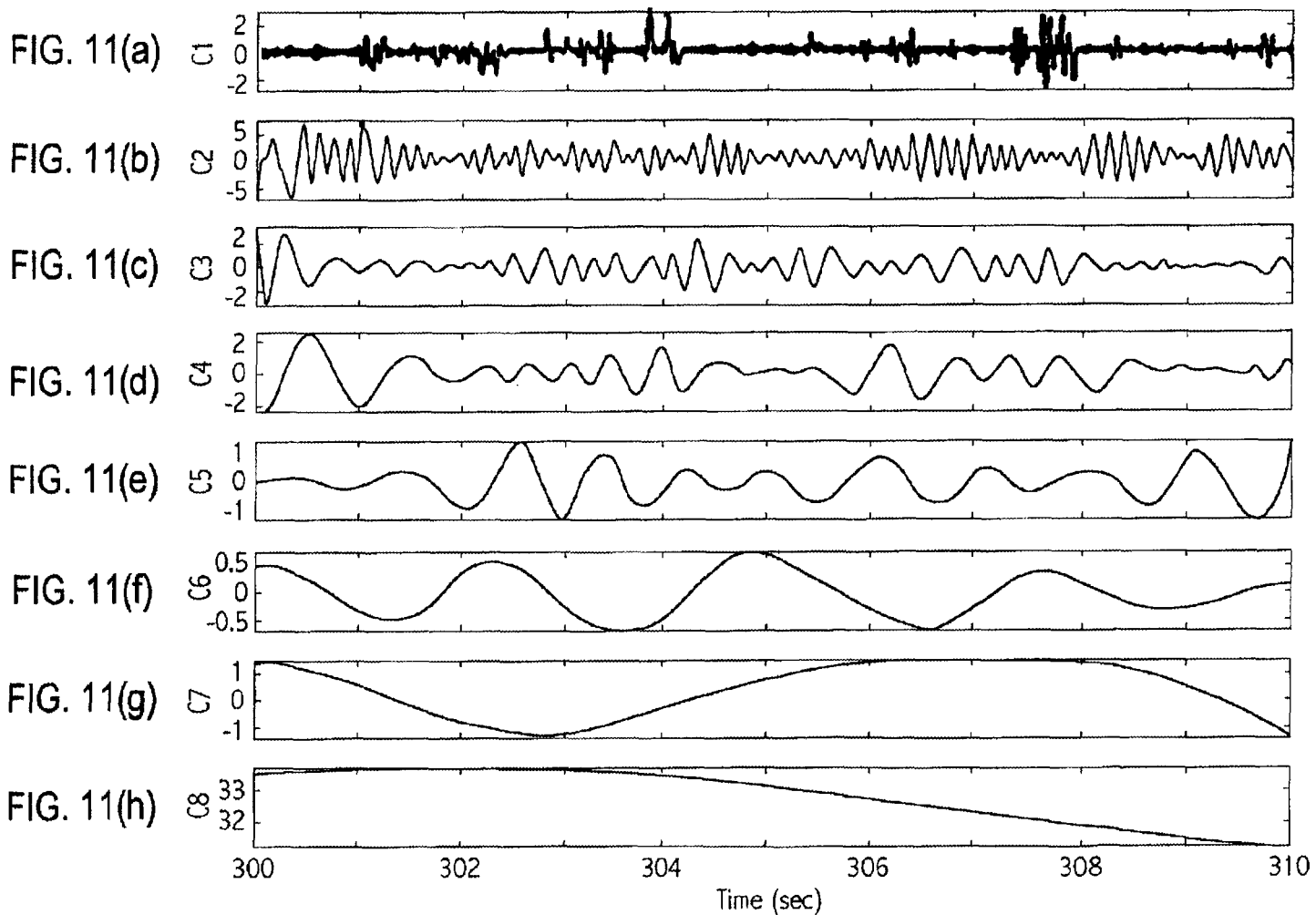
Time (sec)

IMF-COMPONENTS OF THE SECOND 10-SEC. STRIP

FIG. 12(a)

FIG. 12(b)

FIG. 12(c)

FIG. 12(d)

FIG. 12(e)

FIG. 12(f)

FIG. 12(g)

FIG. 12(h)
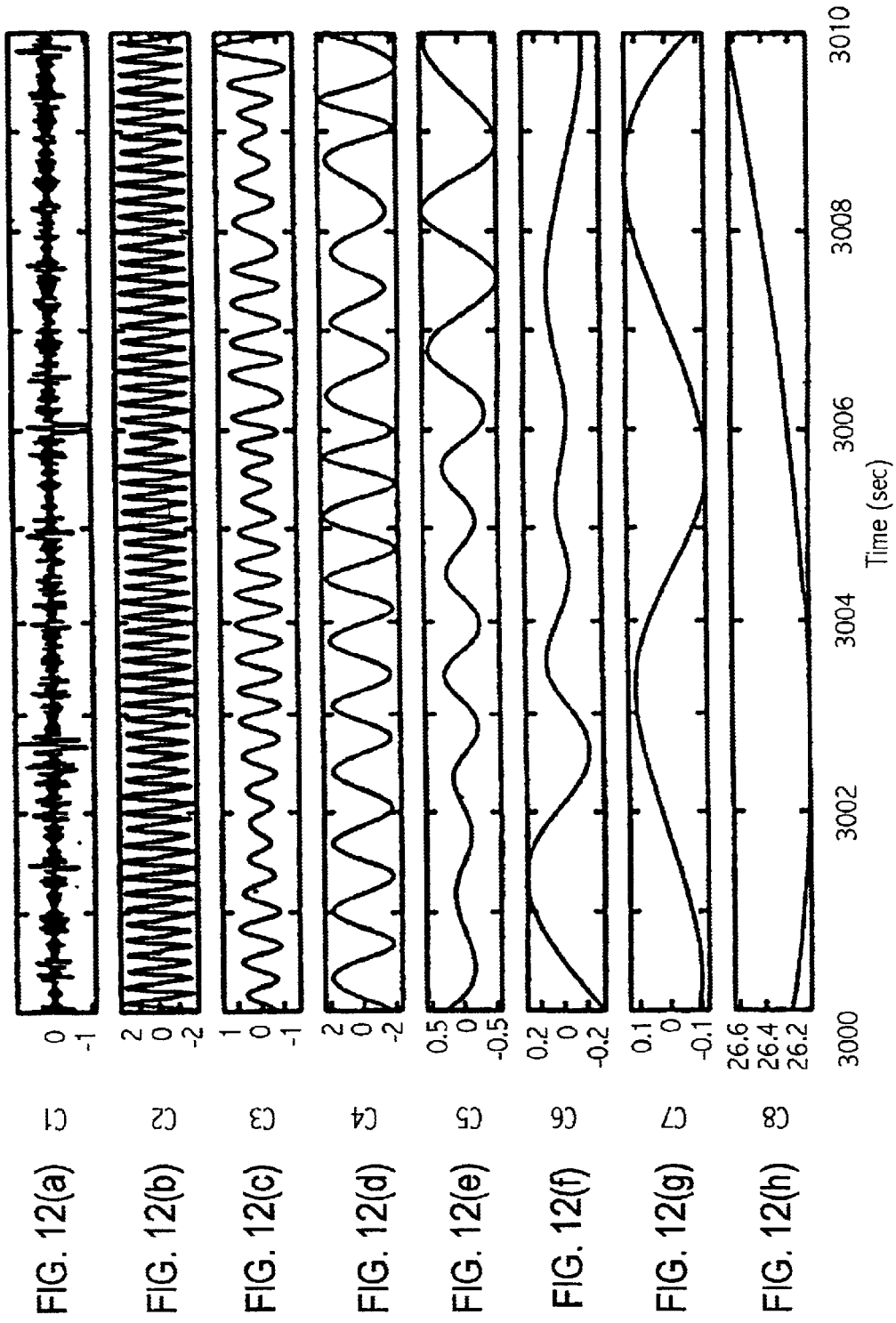
UPPER PANEL: C1+C2+C3; LOWER PANEL: C2+C3+C4
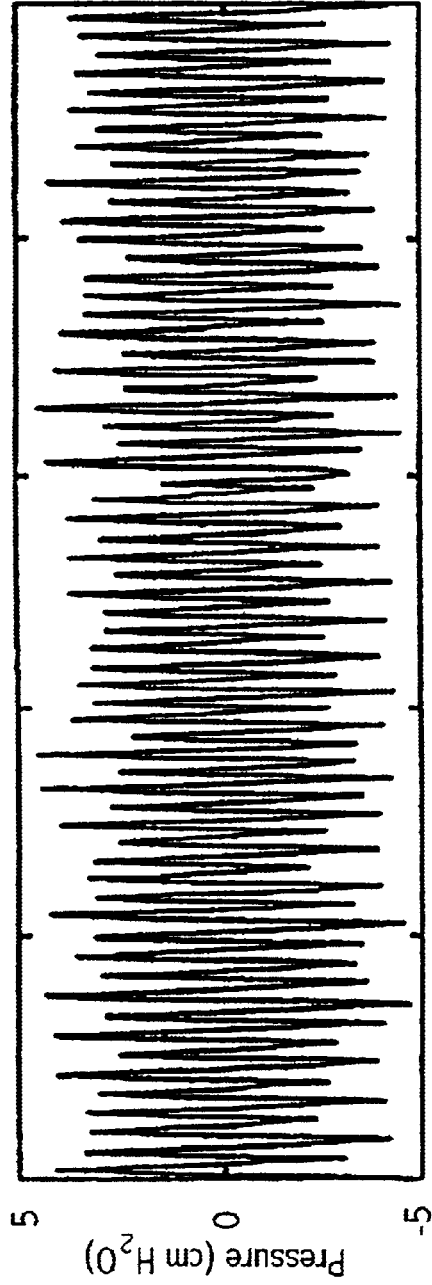
FIG. 13(a)

FIG. 13(b)

FIG. 13(c)
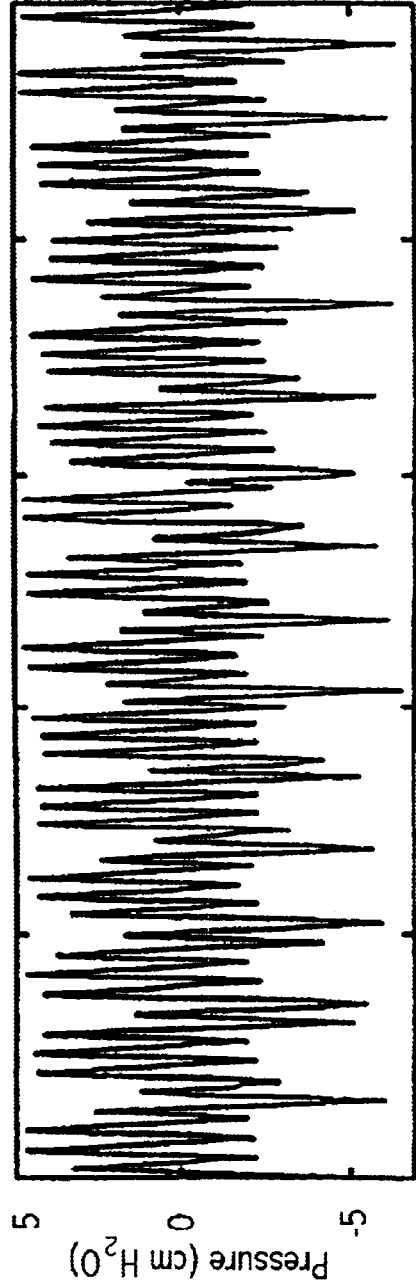
FIG. 14(a)

HHT-SPECTRUM OF THE SECOND 10-SEC STRIP
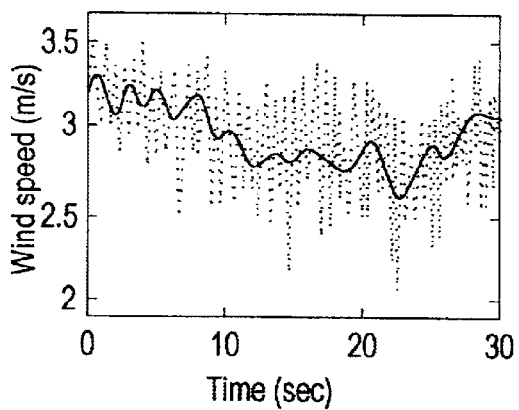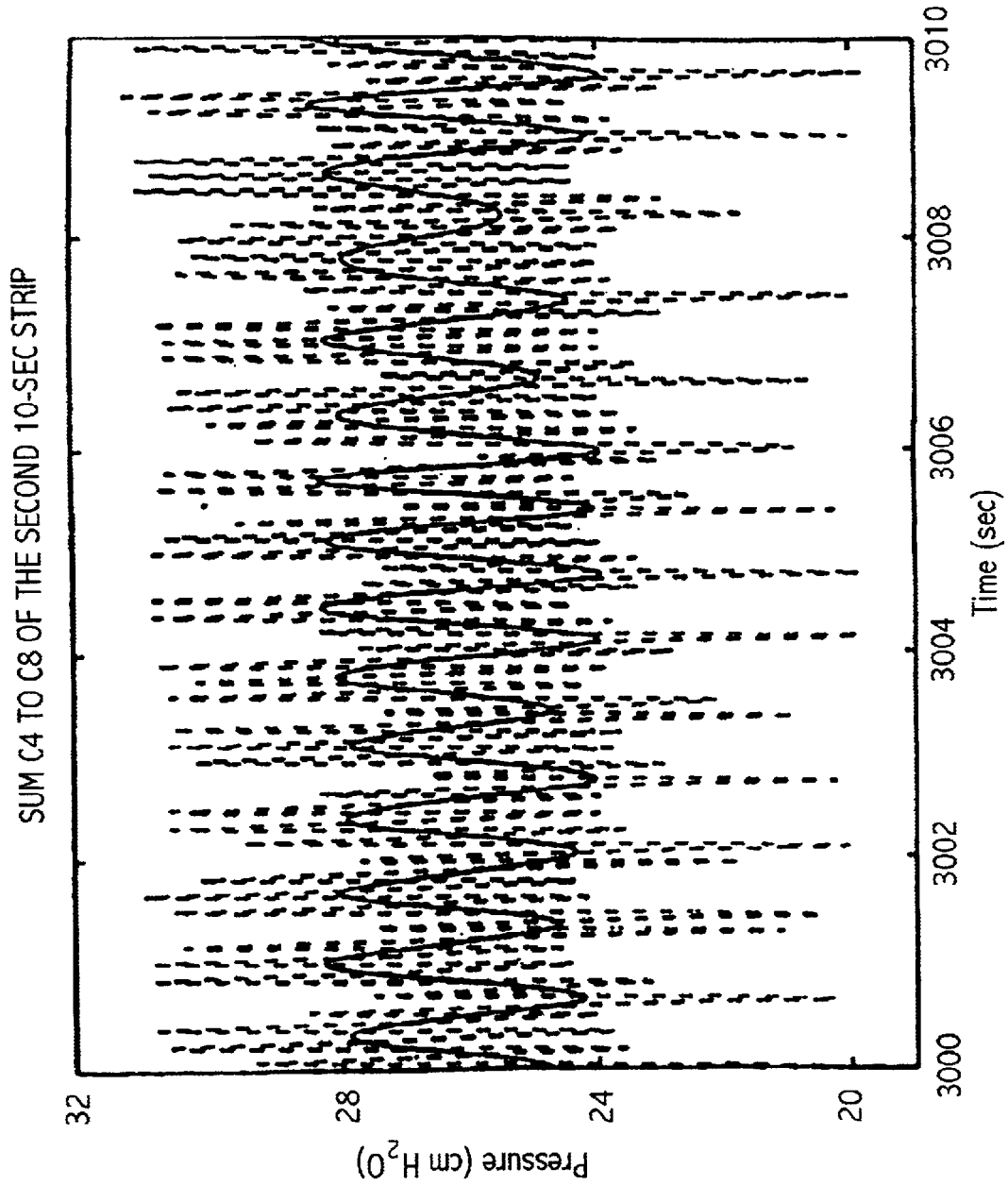
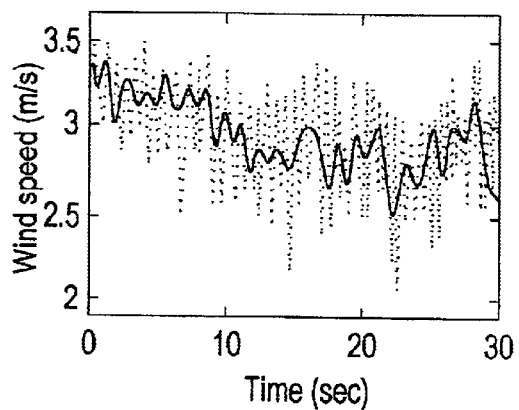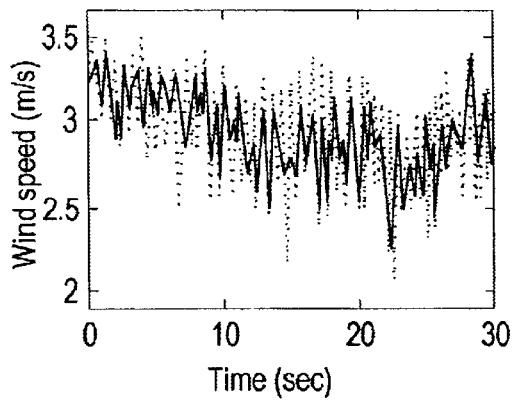HHT-SPECTRUM OF THE FIRST 10-SEC STRIP
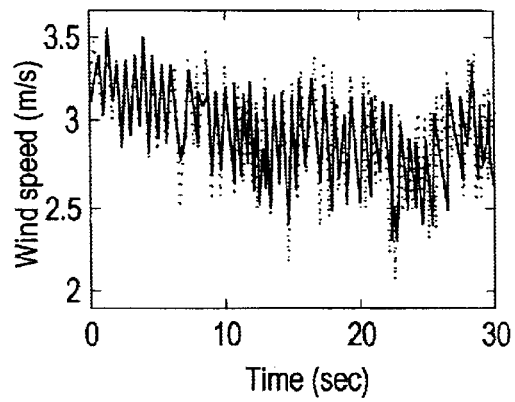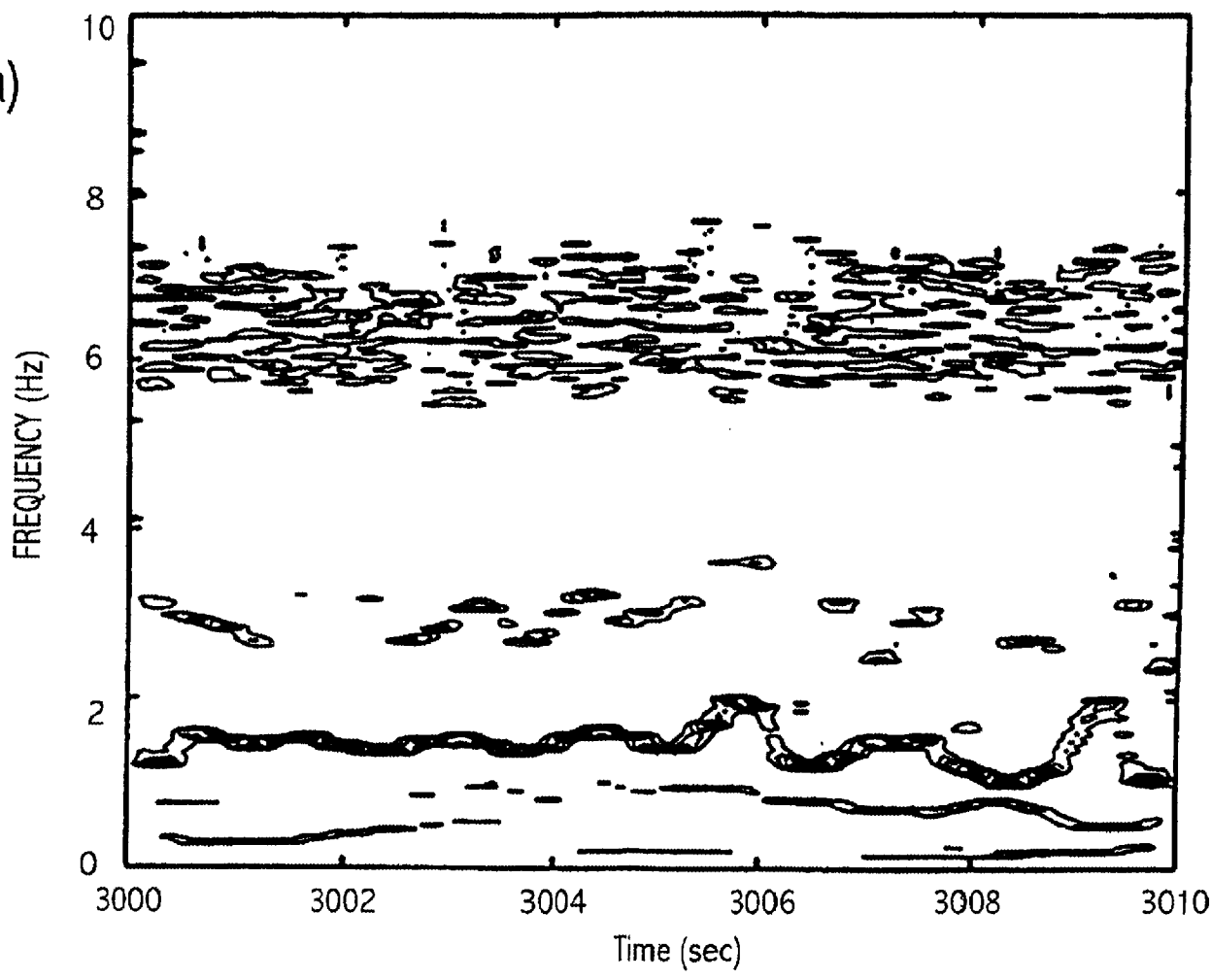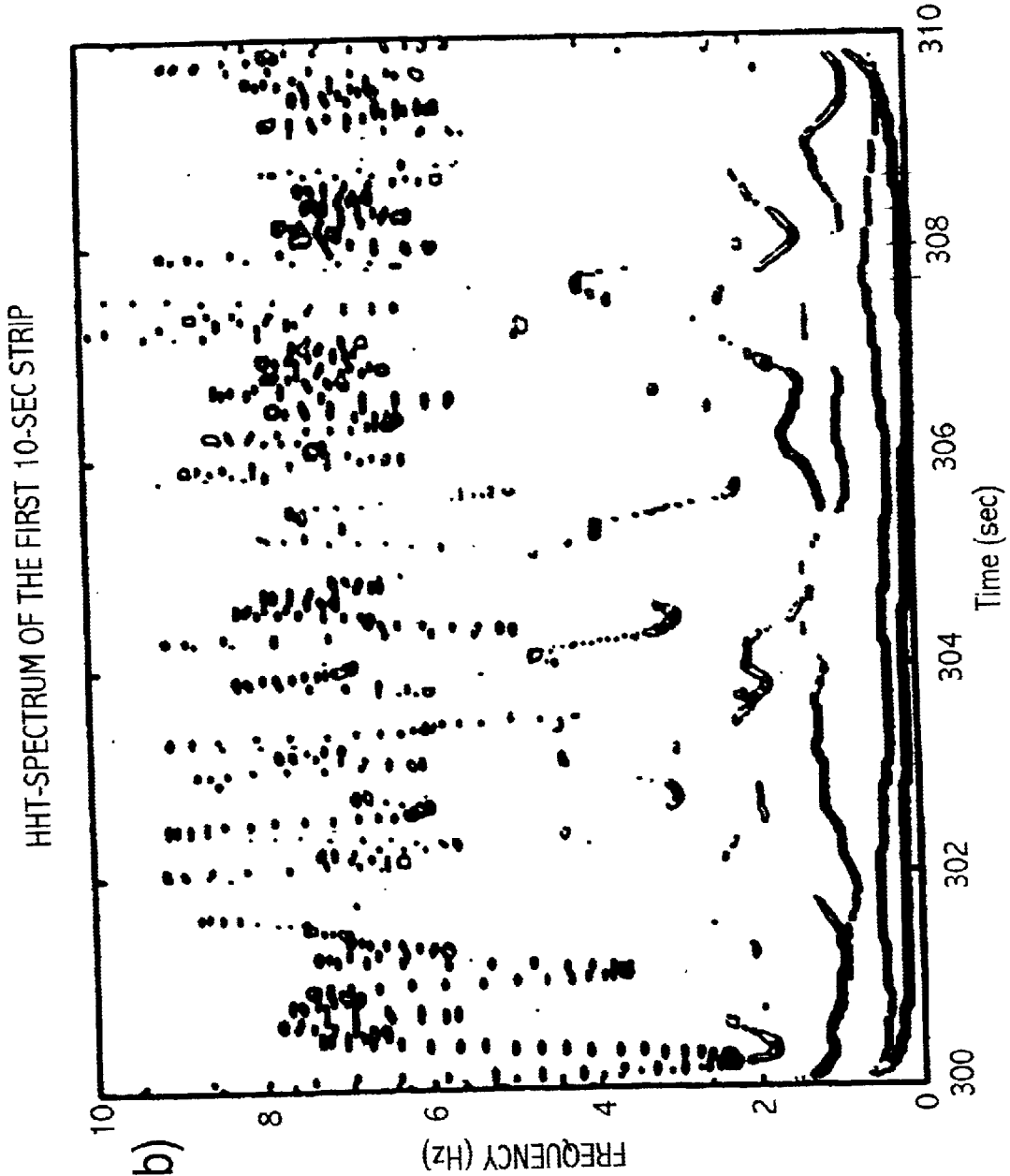
FIG. 14(b)

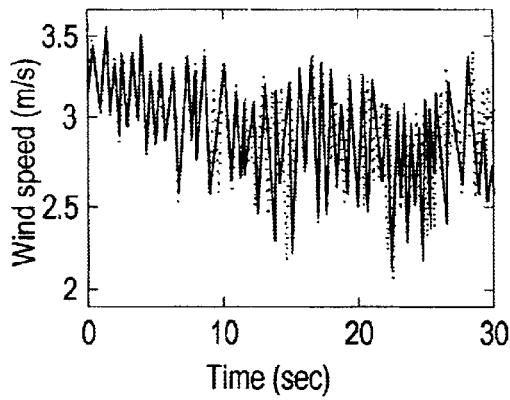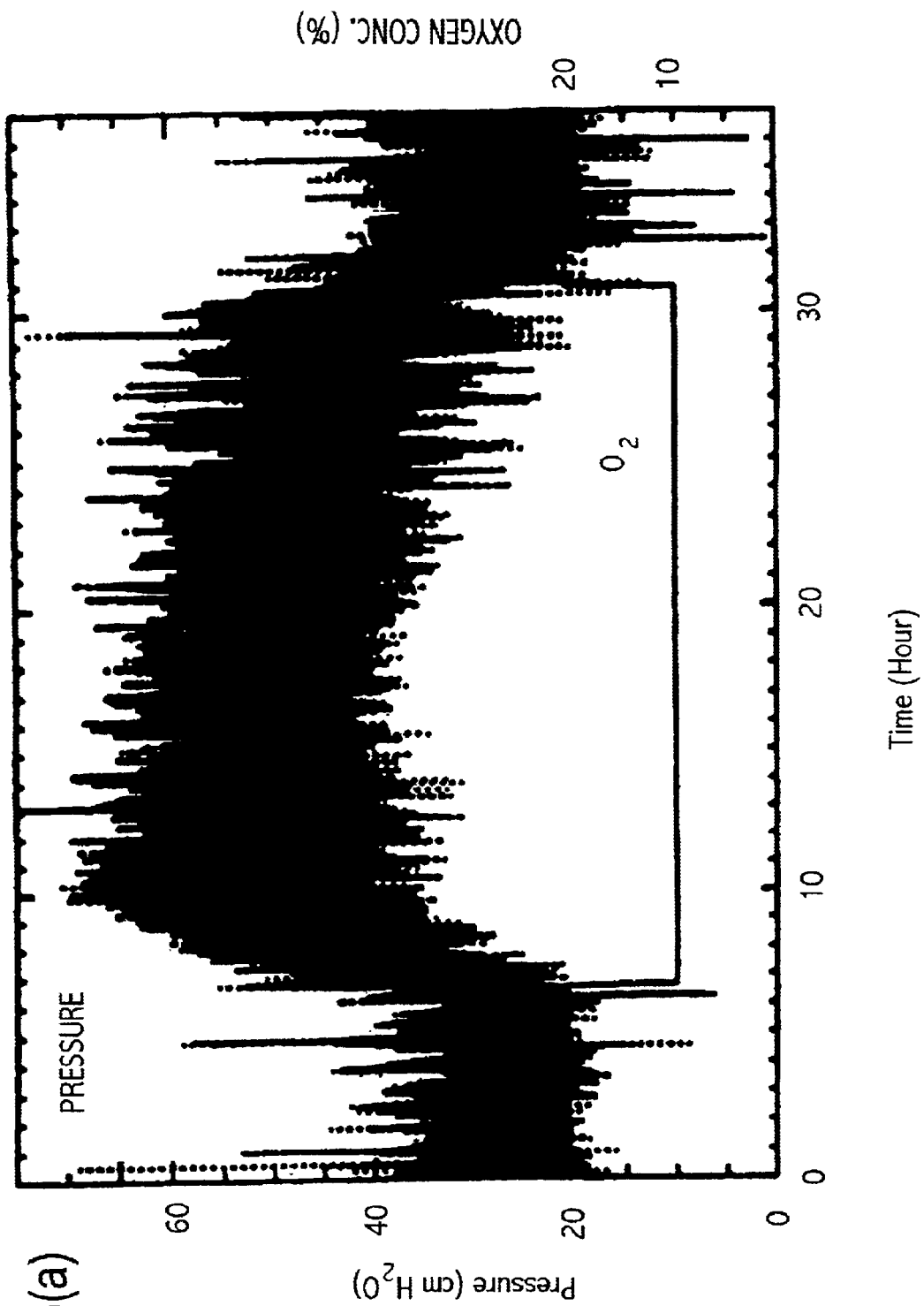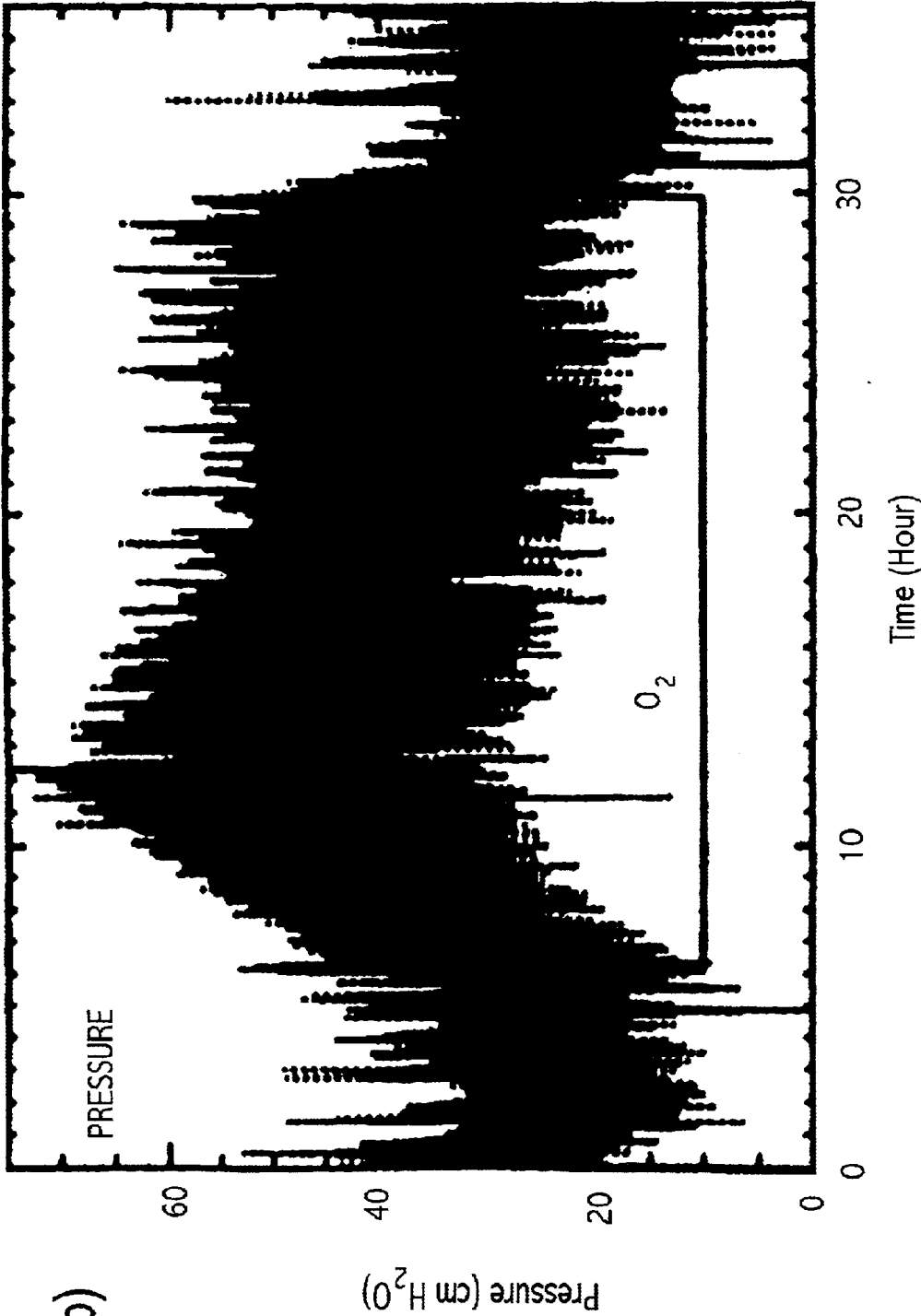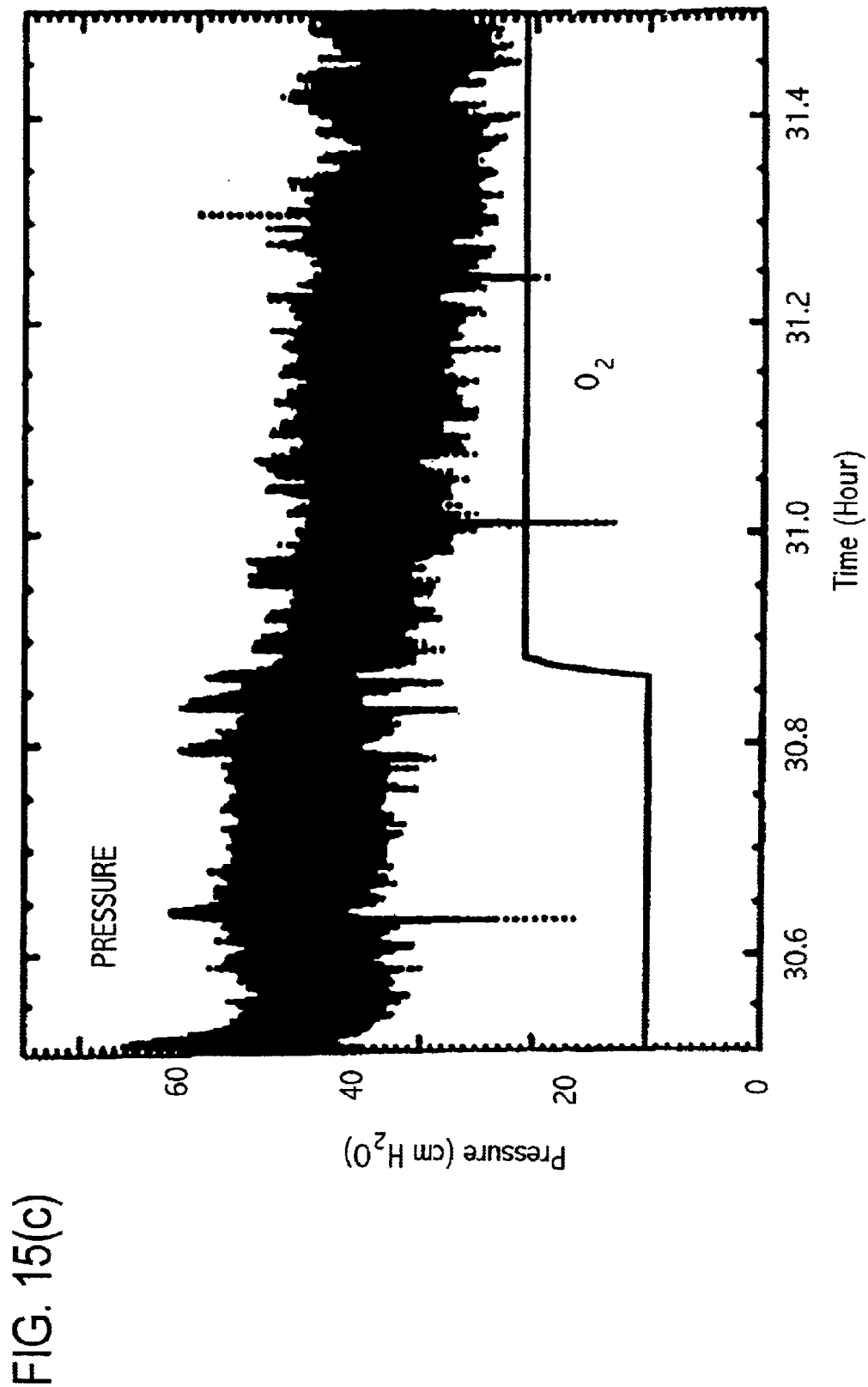FIG. 15(a)

FIG. 15(b)

FIG. 15(c)
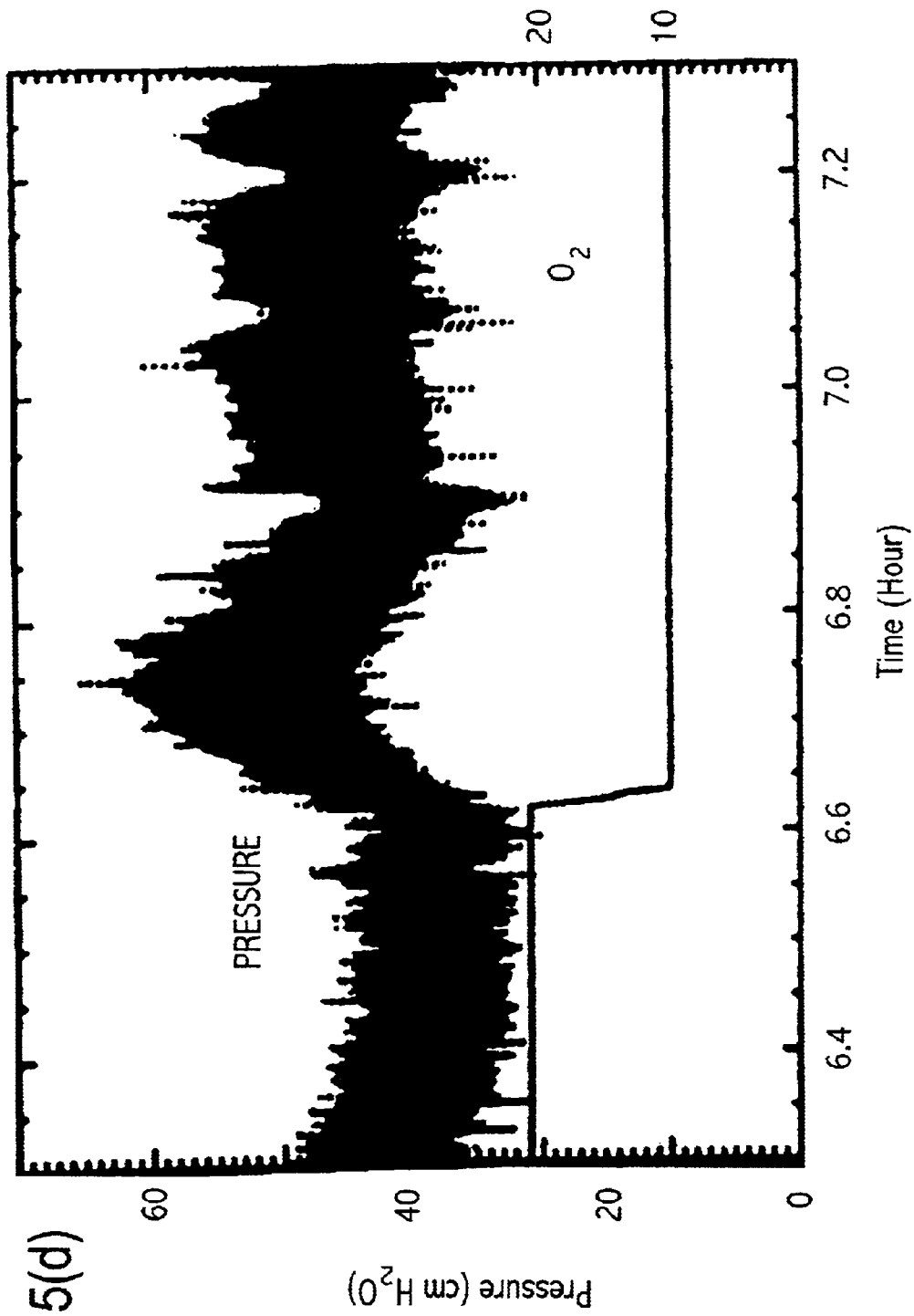
FIG. 15(d)

FIG. 15(e)

FIG. 15(f)

FIG. 15(g)

Blood Pressure (cm H$_2$0)

Time (sec)

FIG. 15(h)

Blood Pressure (cm H₂O)

Time (sec)
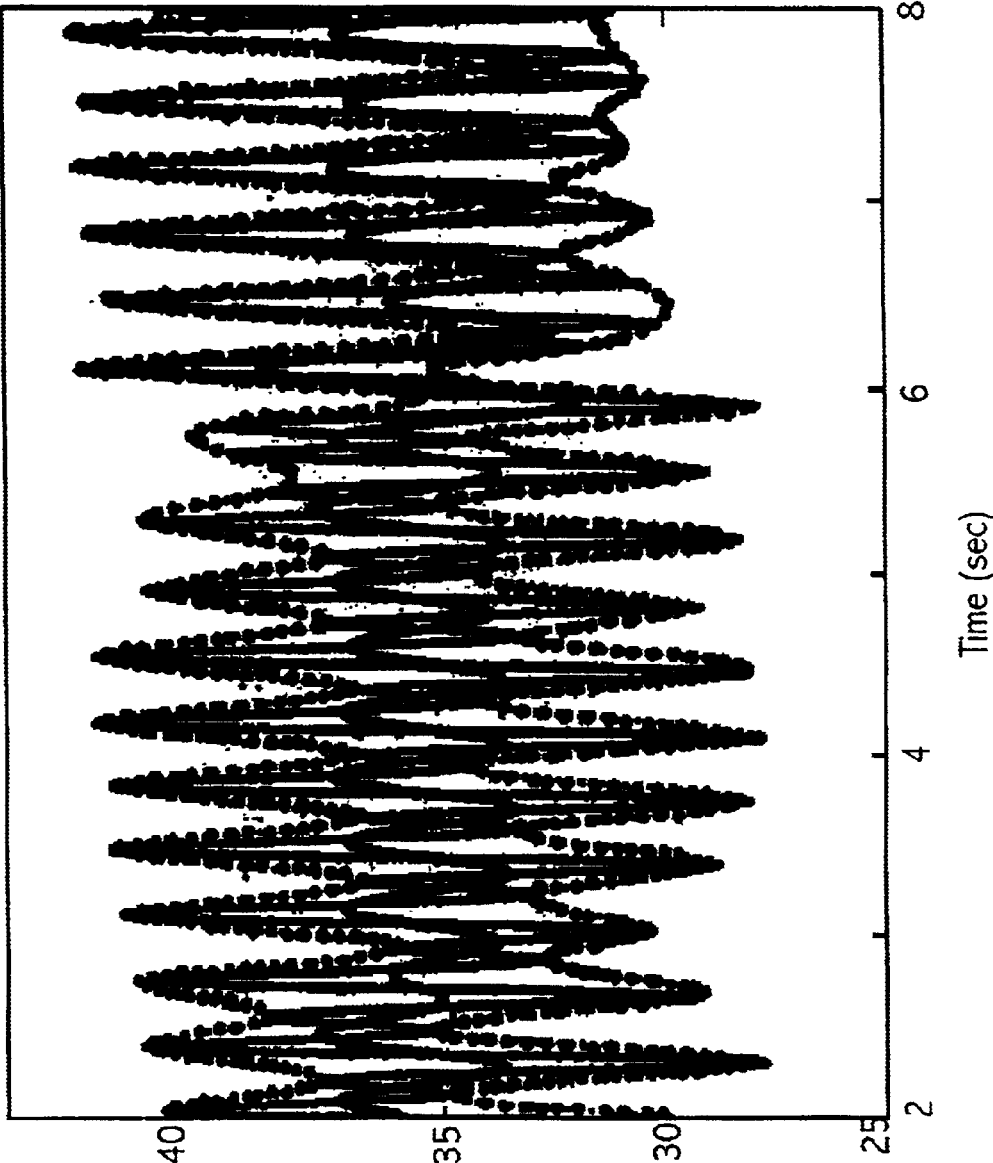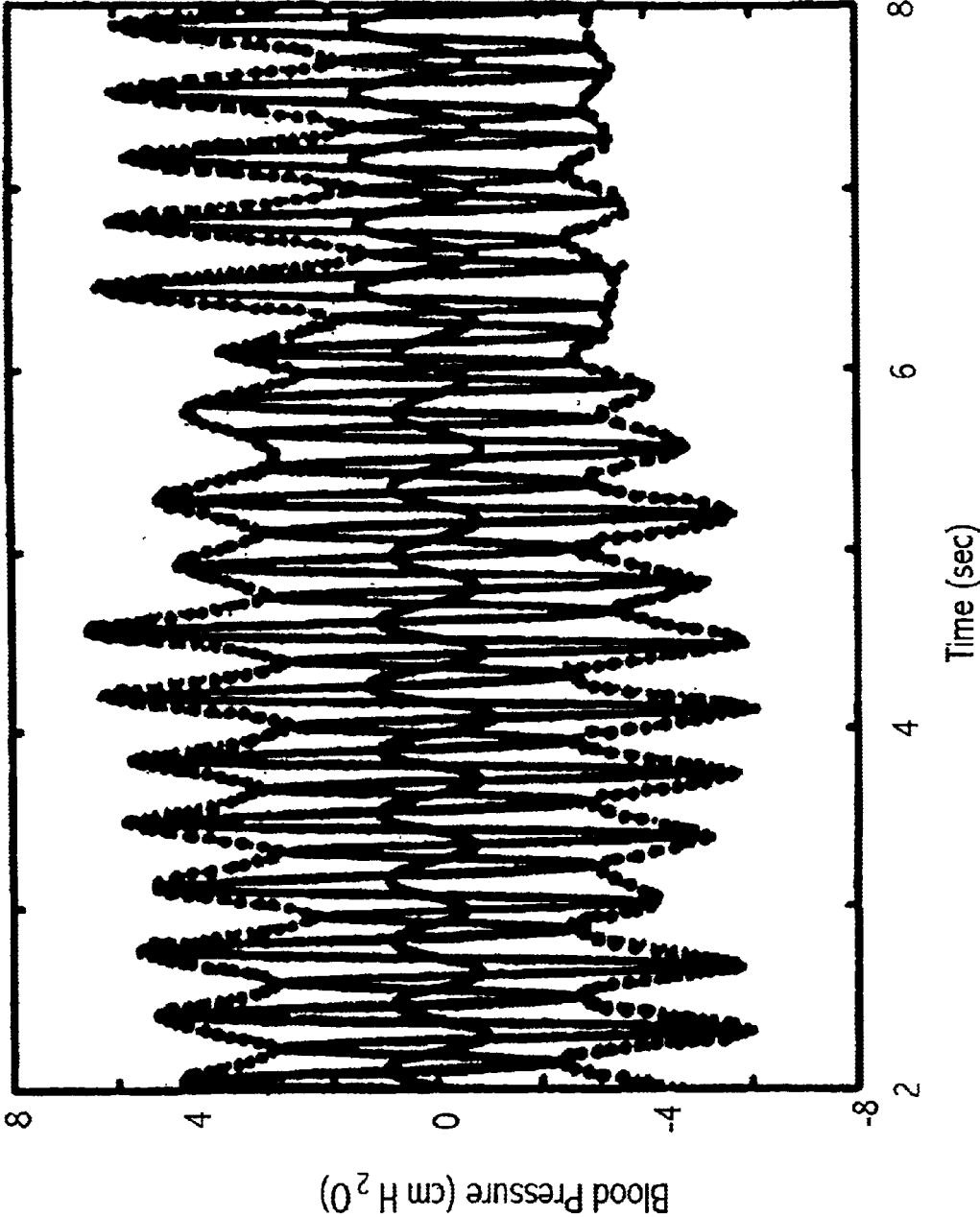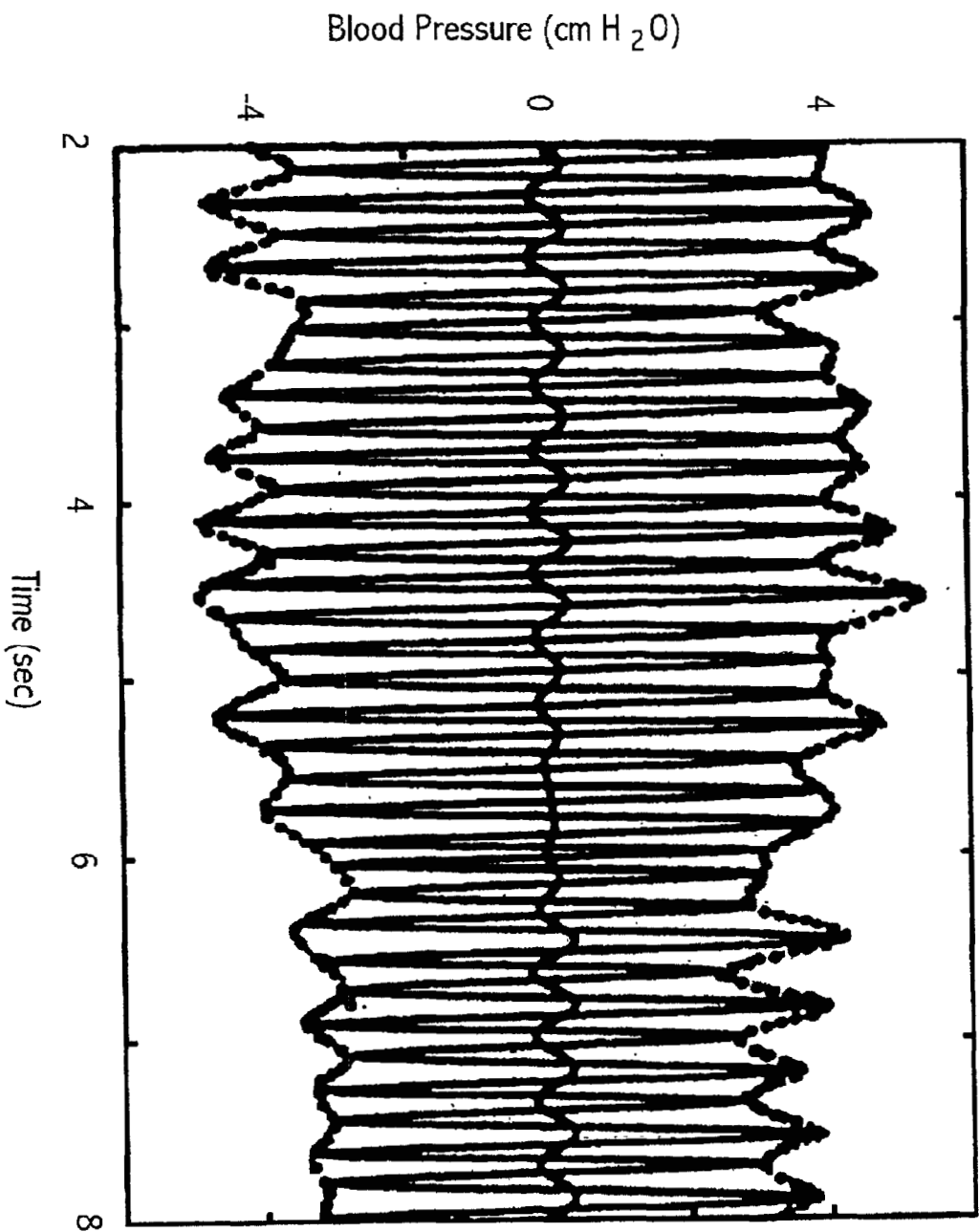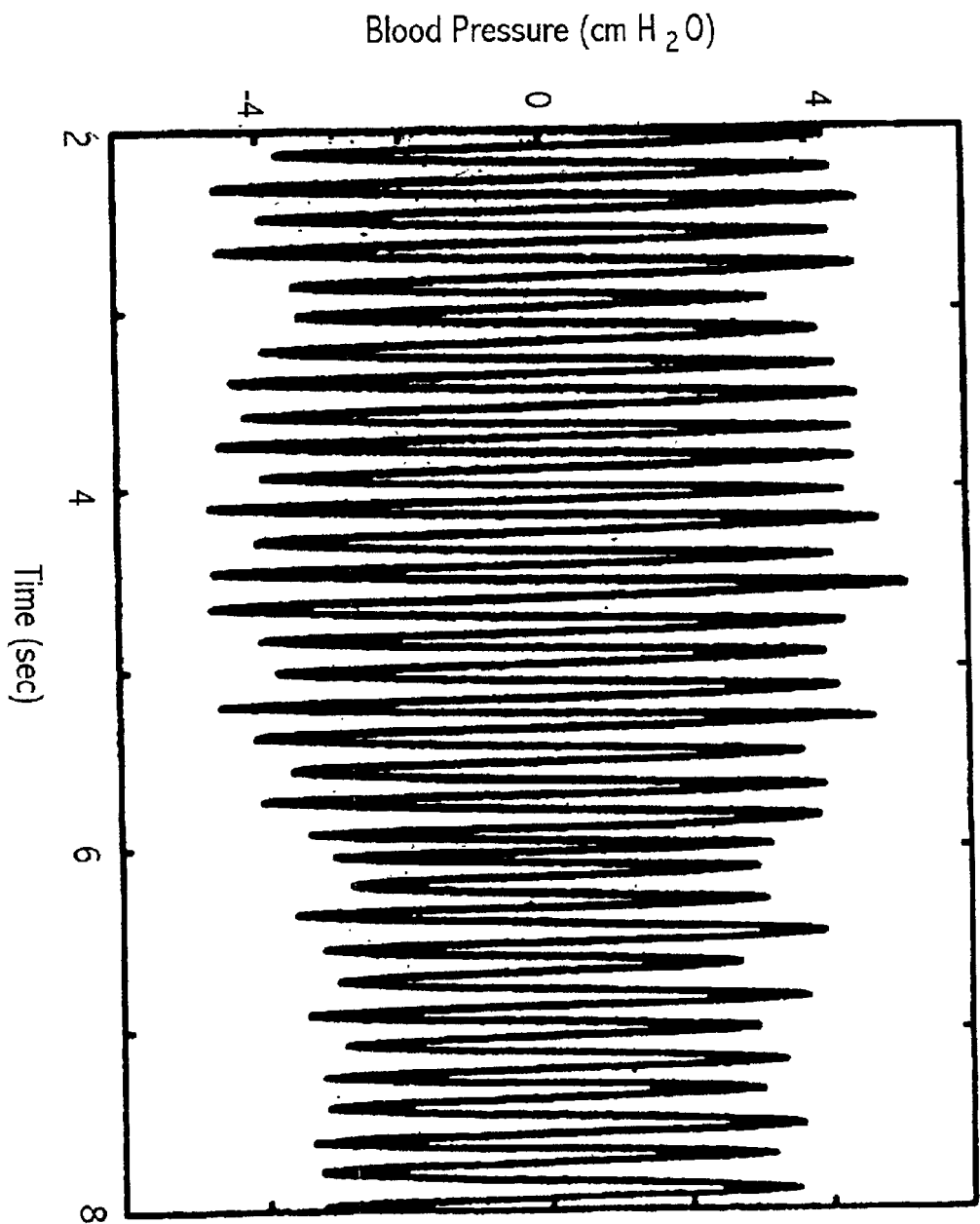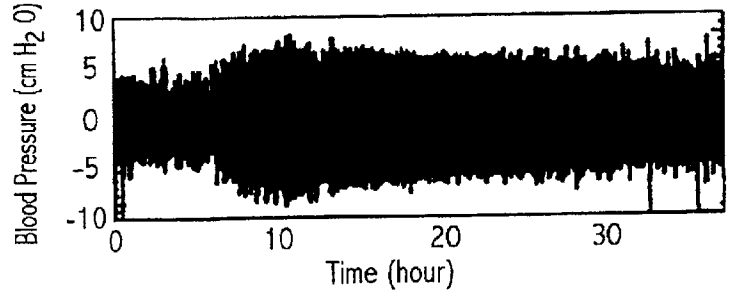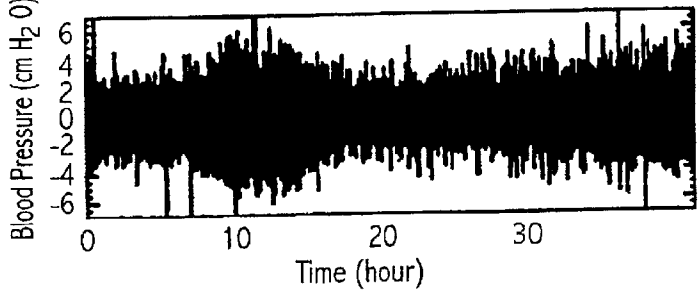
FIG. 16(a)
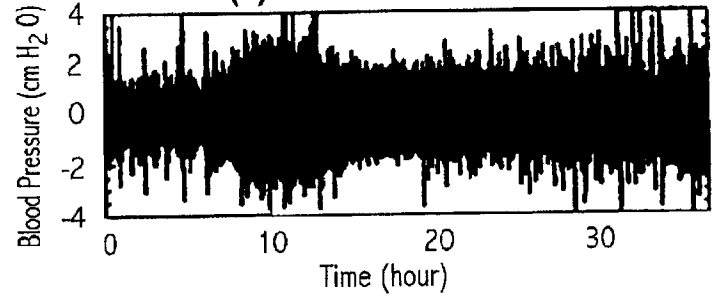
FIG. 16(b)

FIG. 16(c)

FIG. 16(d)

FIG. 16(e)

FIG. 16(f)

FIG. 16(g)

FIG. 16(h)

FIG. 16(i)

FIG. 16(j)

FIG. 16(k)

FIG. 16(l)

FIG. 16(n)

FIG. 16(m)

FIG. 16(p)

FIG. 16(o)

FIG. 17(a)

FIG. 17(b)

FIG. 17(c)

FIG. 17(d)

FIG. 17(e)

FIG. 17(f)

FIG. 18(a)

$$M(t) = 29 + 11(t-t_0)e^{-(t-t_0)/4}$$

$$+15[1-e^{-(t-t_0)/5}]$$

$$\text{for } t > t_0 \quad t_0 = 6.6$$

FIG. 18(b)

$$M(t) = 29 + 11(t-t_0)e^{-(t-t_0)/4}$$

$$+15[1-e^{-(t-t_0)/5}]$$

$$\text{for } t > t_0 \quad t_0 = 6.6$$

FIG. 18(c)

$$M(t) = 29 + 5*60(t-t_0)e^{-(t-t_0)/0.1}$$

$$+11(t-t_0)e^{-(t-t_0)/4} + 15[1-e^{-(t-t_0)/5}]$$

$$\text{for } t > t_0 \quad t_0 = 6.6$$

FIG. 19(a)

$M(t) = 40 + 12[e^{-0.3(t-30.8)} - 1]$

for $t > 30.8$

Blood Pressure (cm H_2O)

Time (hour)

Legend: M(t), $M_{16}(t)$

FIG. 19(b)

$M(t) = 40 + 12[e^{-1.4(t-30.8)} - 1]$ for $t > 30.8$

Blood Pressure (cm H_2O)

Time (hour)

Legend: M(t), $M_{14}(t)$

FIG. 19(c)

for $t > 30.8$

$M(t) = 40 + 12[e^{-1.7(t-30.8)} - 1]$

Blood Pressure (cm H_2O)

Time (hour)

Legend: M(t), $M_{10}(t)$

FIG. 20(a)

FIG. 20(b)

FIG. 20(c)

FIG. 20(d)

FIG. 21

E6(t)

Energy

Time (hour)

FIG. 22(e)

FIG. 22(f)

FIG. 22(g)

FIG. 22(h)

CEK(t): k=13-16

Time (hour)

FIG. 22(a)

FIG. 22(b)

FIG. 22(c)

FIG. 22(d)

CEK(t): k=9-12

Time (hour)

FIG. 23(a)

FIG. 23(b)

FIG. 24

FIG. 25

Heart Beat Rate Data : nn.slp04

FIG. 26

Heart Beat Rate Data : Normal

IMF Normal
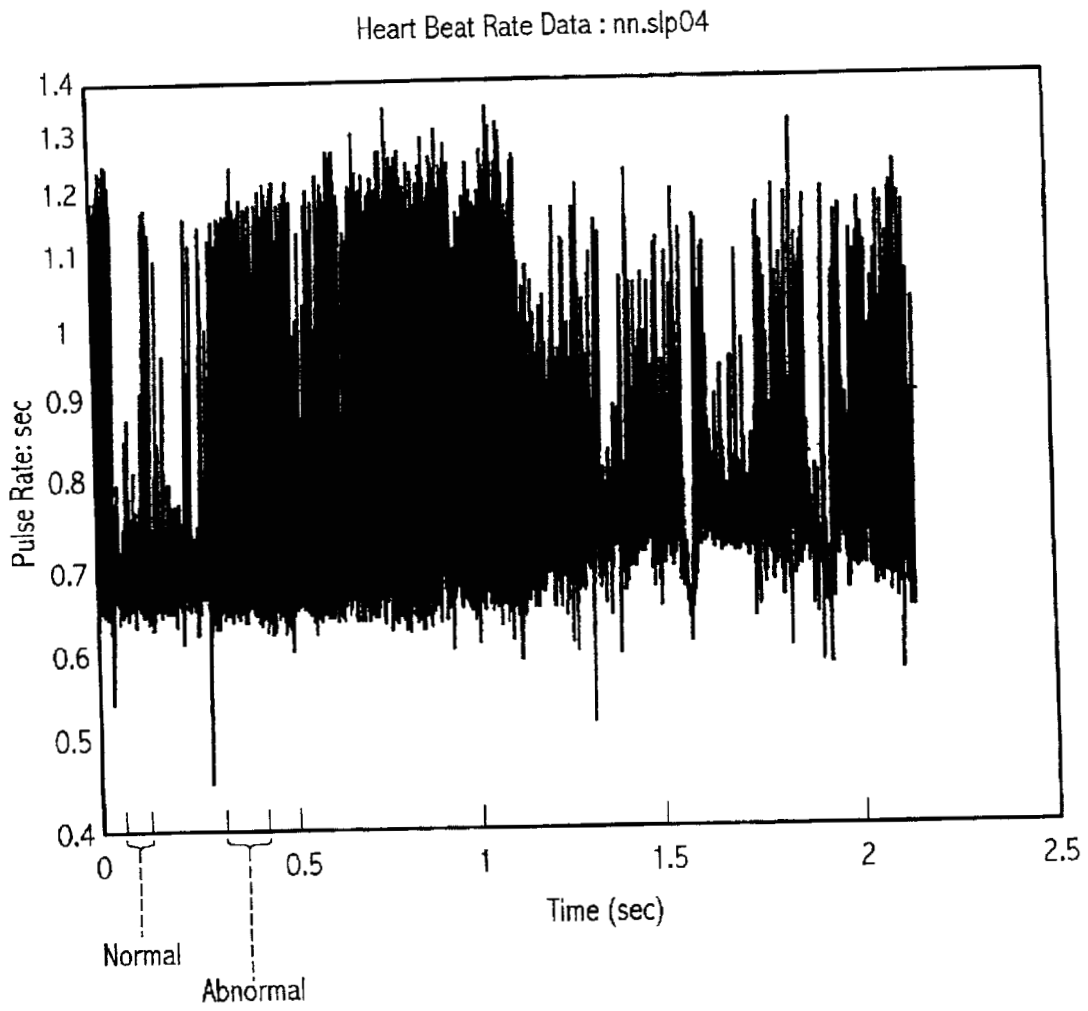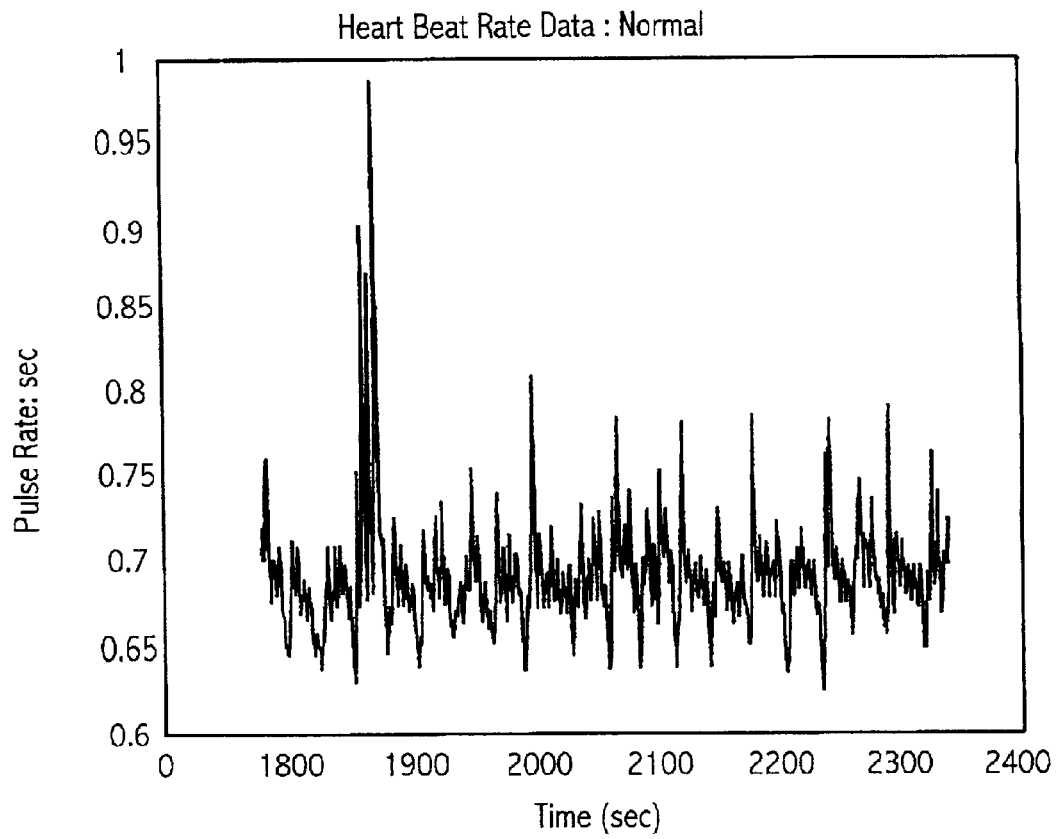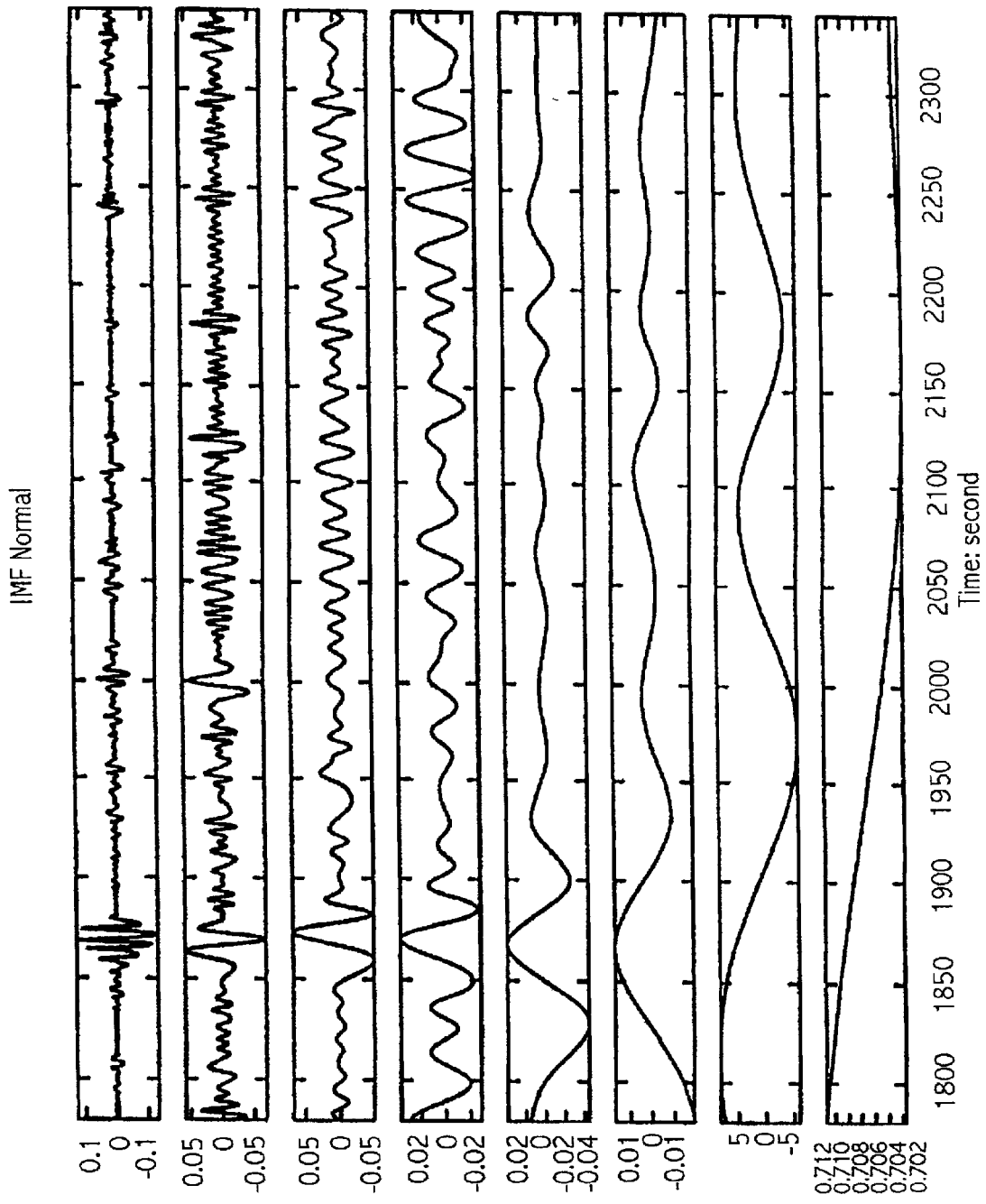
FIG. 27(a)

FIG. 27(b)

FIG. 27(c)

FIG. 27(d)

FIG. 27(e)

FIG. 27(f)

FIG. 27(g)

FIG. 27(h)

Time: second

FIG. 28

Time Series : Normal Section nn.slp04
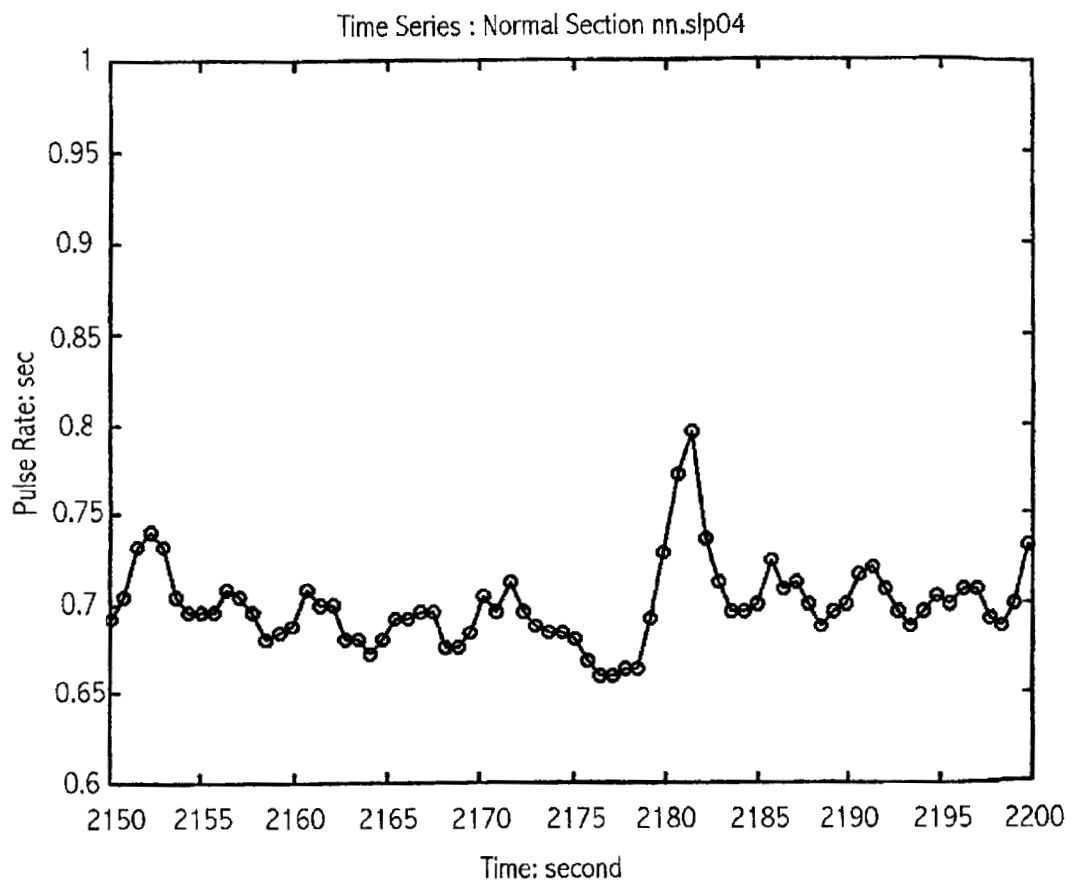
FIG. 29

Hilbert Spectrum : Normal
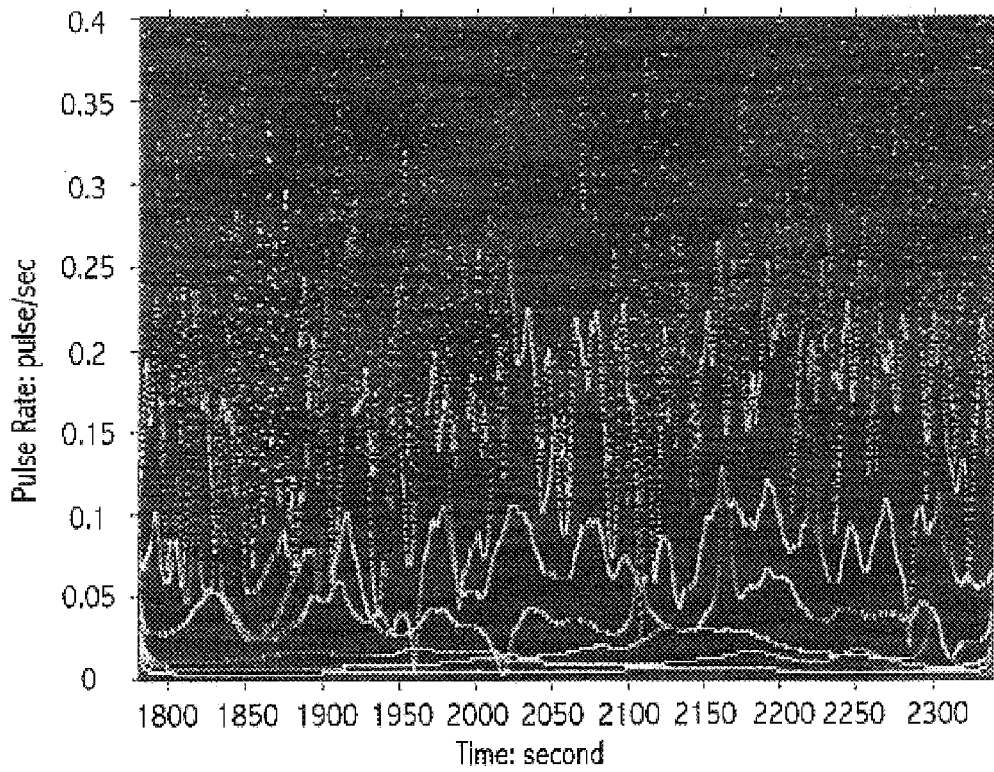
FIG. 30

Heart Beat Rate Data : Abnormal

FIG. 31(a)
FIG. 31(b)
FIG. 31(c)
FIG. 31(d)
FIG. 31(e)
FIG. 31(f)
FIG. 31(g)
FIG. 31(h)
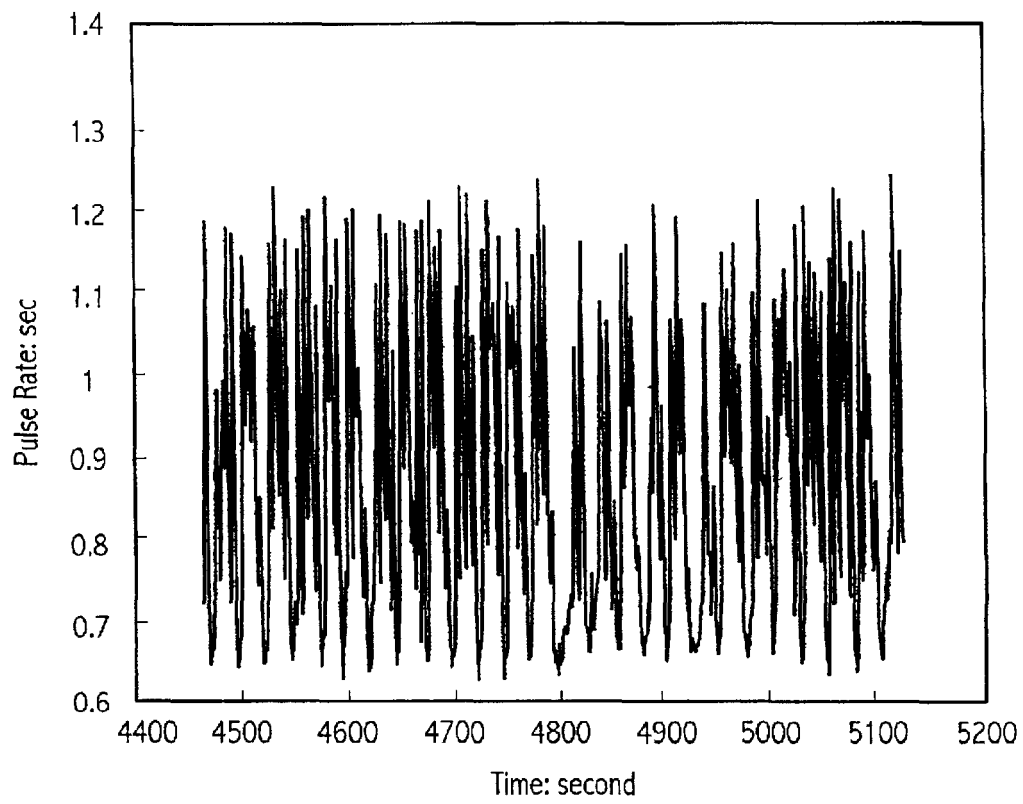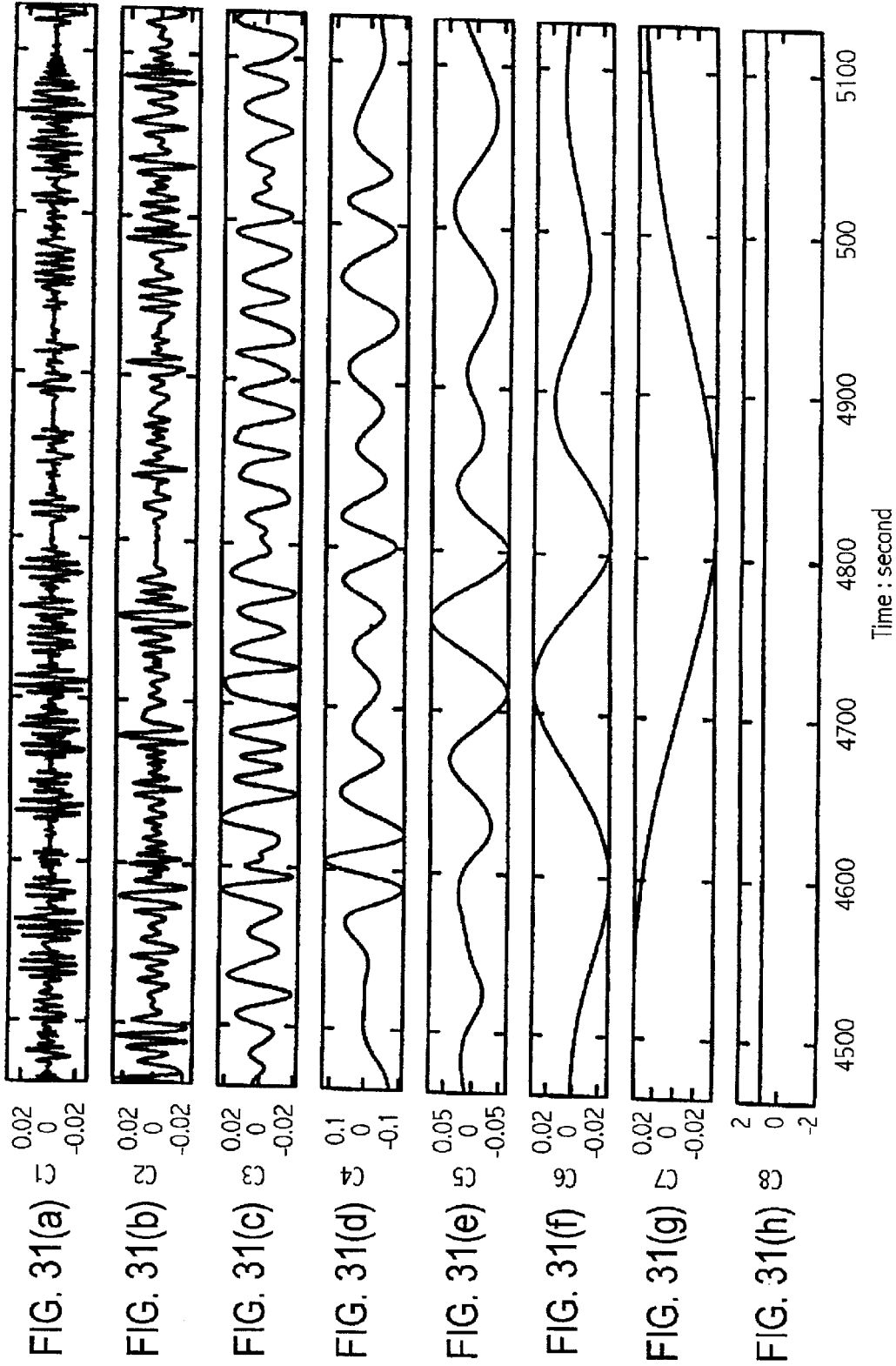
FIG. 32

Time Series : Abnormal Section nn.slp04
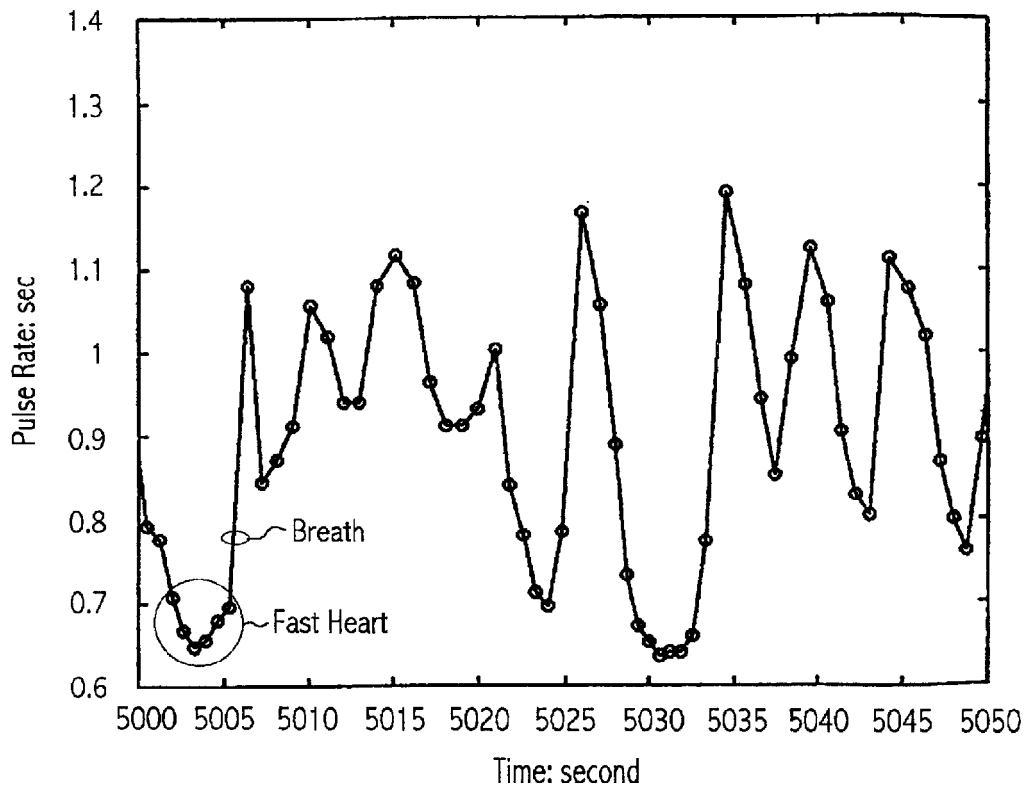
FIG. 33

Hilbert Spectrum

FIG. 34

Seizure Data

FIG. 35(a)   C1

FIG. 35(b)   C2

FIG. 35(c)   C3

FIG. 35(d)   C4

FIG. 35(e)   C5

FIG. 35(f)   C6

FIG. 35(g)   C7

FIG. 35(h)   C8

FIG. 35(i)   C9

Time: minutes
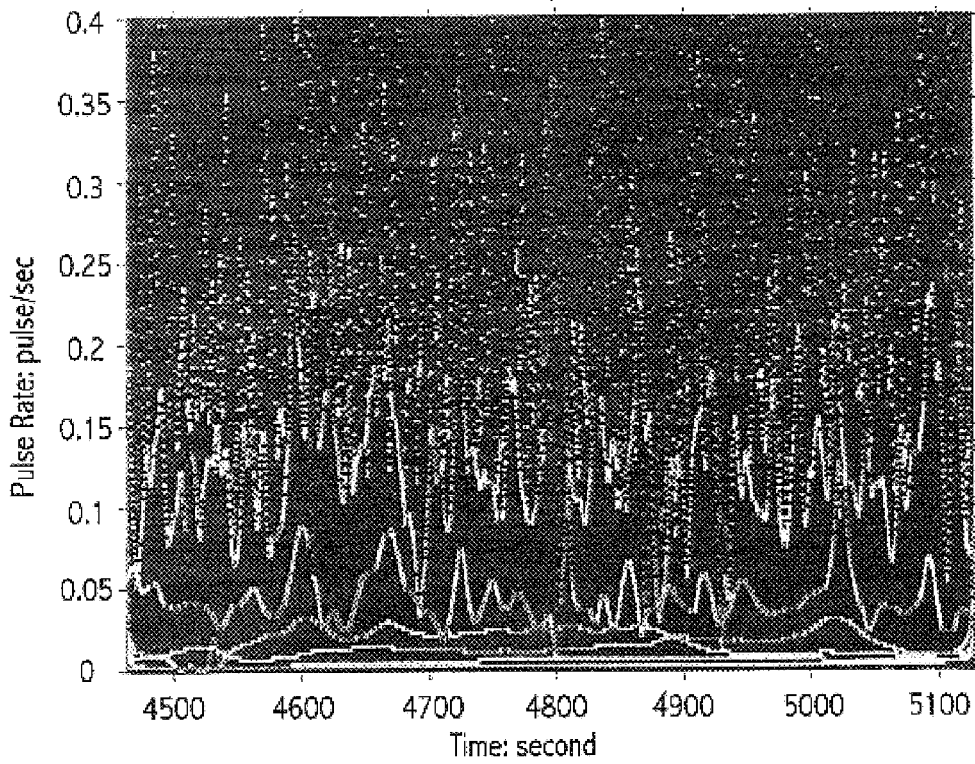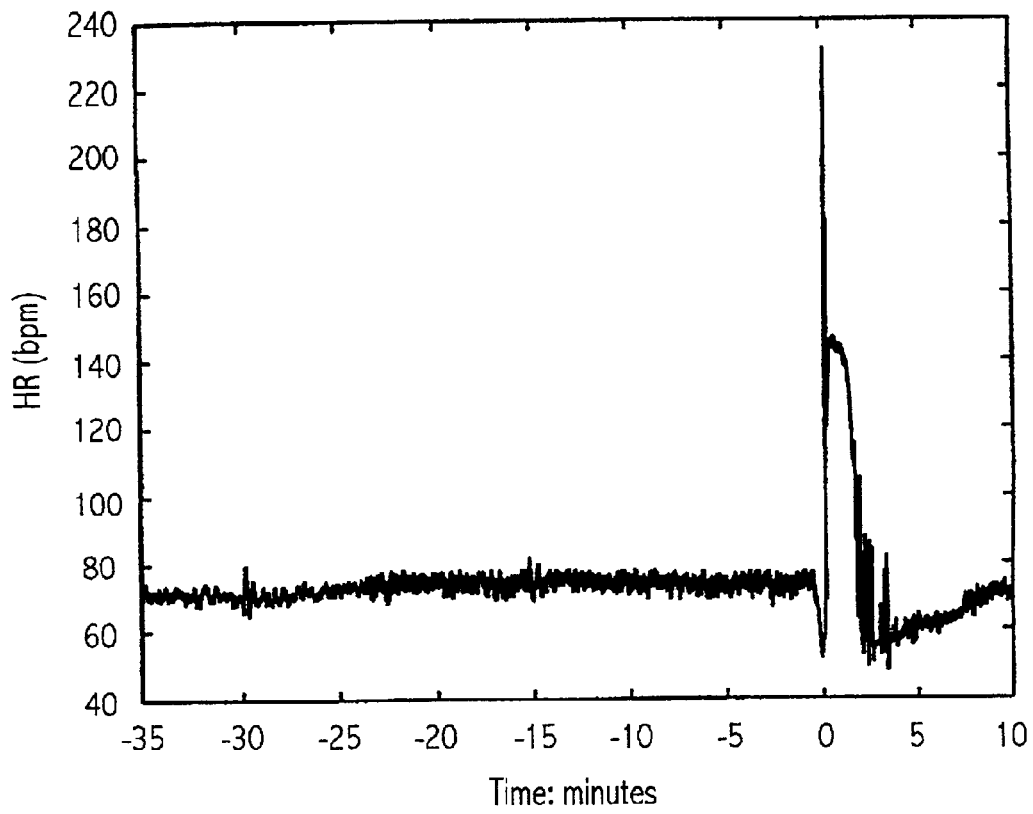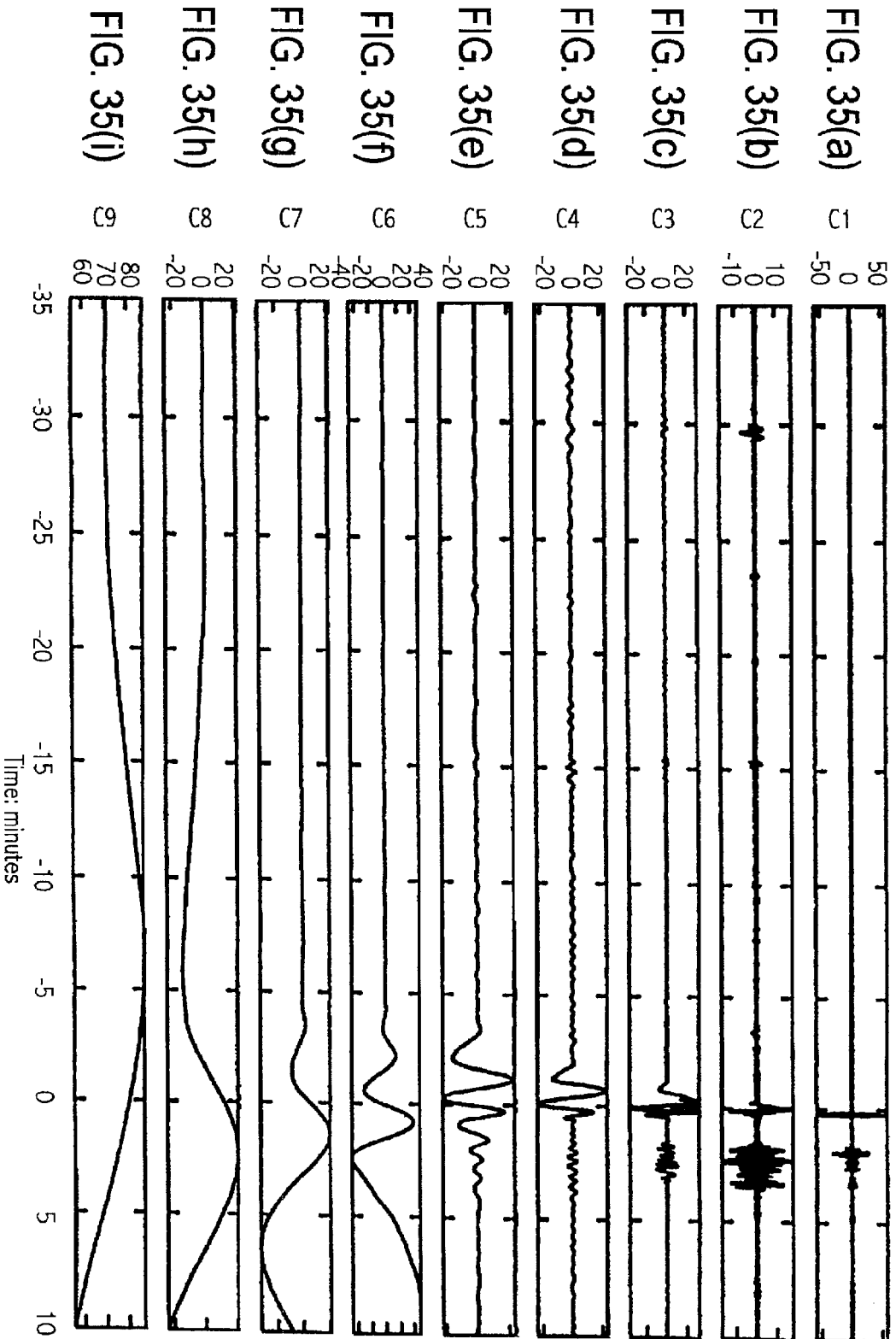
Intrinsic Mode Decomposition

FIG. 36(a)

Time (minutes)

FIG. 36(b)

Time (minutes)

FIG. 36(c)

Time (minutes)
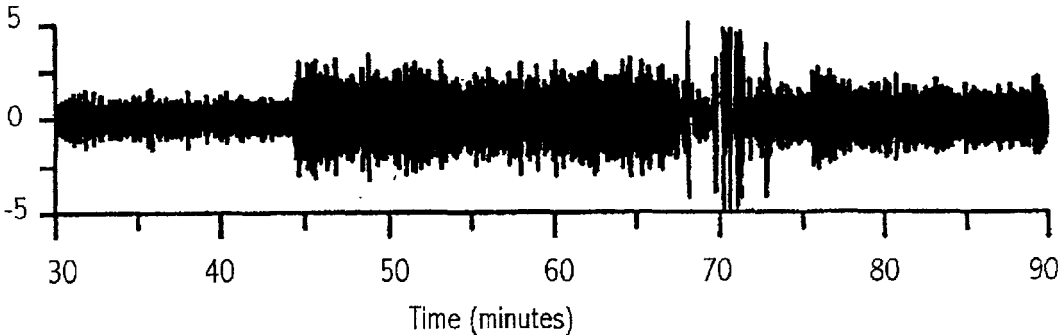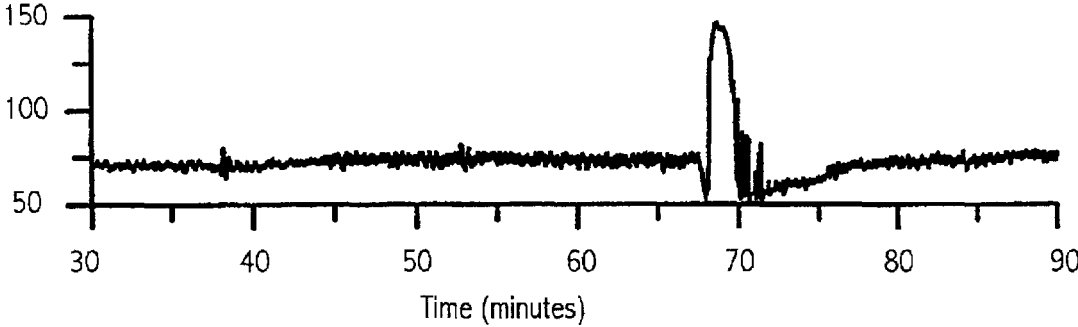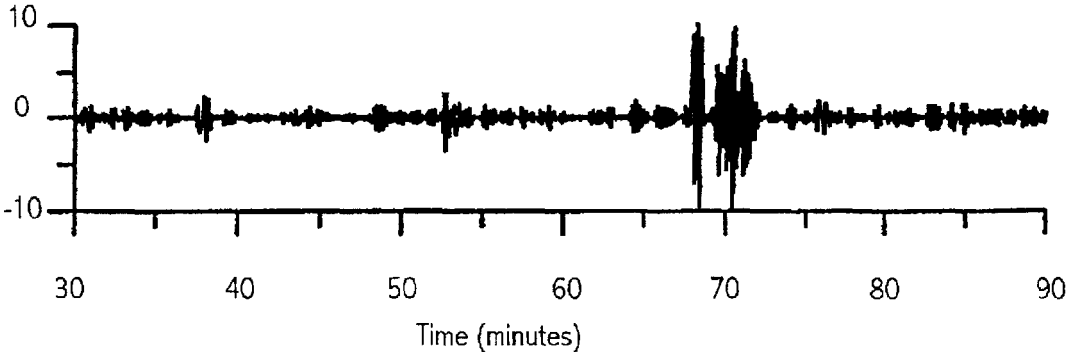
FIG. 37

Hilbert Spectrum Seizure Data

FIG. 38

DATABASE
OF
PHONEMES &
FORMANTS
(HHT BASED)

Extracting
Phonemes &
Formants

HHT

Speech Signal

FIG. 39

TEXT →

| TEXT
Translated
into IMFs | → | Sum IMFs
into
acoustical
signal | → | Play
into
speakers | → |

DATABASE
OF
PHONEMES &
FORMANTS
(HHT BASED)

FIG. 40

FIG. 41

DATABASE
OF
PHONEMES &
FORMANTS
(HHT BASED)

From a Known Speaker

Speaker A

HHT

Extracting
Phonemes &
Formants

HHT Based
Comparison

Verify the
ID of
Speaker

Speaker B

HHT

Extracting
Phonemes &
Formants

HHT Based
Comparison

A = B
or
A ≠ B

FIG. 42

DATABASE
OF
PHONEMES &
FORMANTS
(HHT BASED)

Established Templates Averaged Over
a large number of speeches

Speech Signal

HHT

Extracting
Phonemes &
Formants

Translate
signal into
Text

Grammar
& Syntax
Check

Verified
Text

Extraxtion of
Key Word for
Action

Issue order

Identify Noise
in time

Remove
Noise
IMFs

IMFs &
Hilbert Spectrum

Synthesize
cleaned
signal

EMD

output

Acoustical Signal

FIG. 43

FIG. 44

FIG. 45(a)

Data : Halloon

FIG. 45(b)

Data : Hallooj

FIG. 46(a)

IMF : Halloon

FIG. 46(b)

IMF : Hallooj

FIG. 47(a)

Hilbert Spectrum : Halloon

FIG. 47(b)

Hilbert Spectrum : Hallooj

FIG. 48(a)

Fourier Spectrogram (1024): Balloon

FIG. 48(b)

Fourier Spectrogram (1024): Hallooj

FIG. 49(a)

Detailed Hilbert Spectrum : Halloon

FIG. 49(b)

Detailed Hilbert Spectrum ; Hallooj

FIG. 50

Fourier Spectrogram (128) : Halloon

FIG. 51(a)

Detailed Fourier Spectrogram : Halloon

FIG. 51(b)

Detailed Fourier Spectrogram (128) : Halloon

FIG. 52(a)

Detailed Morlet Wavelet Spectrum : Halloon

FIG. 52(b)

Detailed Morlet Wavelet Spectrum : Halloo

FIG. 53(a)

Detailed Hilbert Spectrum : Halloonx100

FIG. 53(b)

Detailed Hilbert Spectrum : Halloonx200

FIG. 54(a)

Detailed Hilbert Spectrum and Data : Halloon

FIG. 54(b)

Detailed Hilbert Spectrum and Data : Halloon

FIG. 55

Data : Halloo+Ding, Halloo, and Ding

FIG. 56(a)

IMF : Halloo & Ding

FIG. 56(b)

IMF : Halloo & Ding

FIG. 57(a)

IMF : Filtered Halloo & Ding

FIG. 57(b)

IMF : Ding

Time : second*44100

FIG. 58(a)

Hilbert Spectrum : Halloo+ding

FIG. 58(b)

Hilbert Spectrum : HHT filtered Halloo+ding

FIG. 59(a)

Data : HHT and Fourier filtered "Halloo+ding"

FIG. 59(b)

Fourier Spectra for the Halloo+Ding and Various Filtered Data

FIG. 60(a)

Data : TR095

FIG. 60(b)

Data : TR120

FIG. 60(c)

Data : TR200

FIG. 61(a)

IMF TR095

FIG. 61(b)

IMF TR120

Time : second

FIG. 61(c)

IMF TR200

FIG. 62(a)

Hilbert Spectrum : TR095

FIG. 62(b)

FIG. 62(c)

1

EMPIRICAL MODE DECOMPOSITION FOR ANALYZING ACOUSTICAL SIGNALS

CROSS-REFERENCES TO RELATED APPLICATIONS

This Application is a Continuation-In-Part of application Ser. No. 10/011,206, "Empirical Mode Decomposition Apparatus, Method and Article of Manufacture for Analyzing Biological Signal and Performing Curve Pitting," filed on Dec. 10, 2001, which claims priority under U.S.C. §120 to parent application Ser. No. 09/282,424 filed on Mar. 31, 1999 and now U.S. Pat. No. 6,381,559 which claims priority under U.S.C. §120 to parent application Ser. No. 08/872,586 filed on Jun. 10, 1997 and now U.S. Pat. No. 5,983,162, which claims priority under 35 U.S.C. §119(e) to U.S. Provisional application Ser. No. 60/023,411 filed on Aug. 14, 1996 and Ser. No. 60/023,822 filed on Aug. 12, 1996. This application claims priority under 35 U.S.C. §119(e) to U.S. Provisional application Ser. No. 60/269,422 filed on Feb. 14, 2001. All of the above provisional and non-provisional patent applications are hereby incorporated by reference.

ORIGIN OF INVENTION

The inventor of the invention described herein is an employee of the United States Government. Therefore, the invention may be manufactured and used by or for the Government for governmental purposes without the payment of any royalties thereon or therefor.

COPYRIGHT NOTIFICATION

Portions of this patent application contain materials that are subject to copyright protection. The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent disclosure, as it appears in the Patent and Trademark Office patent file or records, but otherwise reserves all copyright rights whatsoever.

BACKGROUND OF THE INVENTION

This invention generally relates to a signal analysis method. The results of processing several examples of biological and acoustical signals are discussed herein to show the particular utility of the invention in that field and to further demonstrate the broad applicability of the invention.

Although the present invention finds utility in processing acoustical signals, it is to be understood that any signal representative of a real world phenomenon such as a signal representative of a physical process including electrical, mechanical, biological, acoustical, chemical, optical, geophysical or other process(es) may be analyzed and thereby more fully understood by applying the invention thereto. The real world signals to which the invention finds utility include a wide variety of real world phenomena such as the behavior of a stock market, population growth, traffic flow, etc. Furthermore, the term "real world signal" also includes "physical signals" representative of physical processes such as the electrical, mechanical, biological, chemical, acoustical, optical, geophysical process(es) mentioned above.

Although the invention is not limited to a particular type of signal processing and includes the full range of real world data representative of processes or phenomena or combinations thereof, it is most useful when such real world signals are nonlinear and nonstationary.

2

DESCRIPTION OF RELATED ART

In the parent application, several examples of geophysical data signals representative of earthquakes, ocean waves, tsunamis, ocean surface elevation and wind were processed to show the invention's wide utility to a broad variety of signal types. The techniques disclosed therein and elaborated upon herein represent major advances in physical signal processing.

Previously, analyzing signals, particularly those having nonlinear and/or nonstationary properties, was a difficult problem confronting many industries. These industries have harnessed various computer implemented methods to process data signals measured or otherwise taken from various processes such as electrical, mechanical, optical, biological, and chemical processes. Unfortunately, previous methods have not yielded results which are physically meaningful.

Among the difficulties found in conventional systems is that representing physical processes with physical signals may present one or more of the following problems:

(a) The total data span is too short;

(b) The data are nonstationary; and

(c) The data represent nonlinear processes.

Although problems (a)–(c) are separate issues, the first two problems are related because a data section shorter than the longest time scale of a stationary process can appear to be nonstationary. Because many physical events are transient, the data representative of those events are nonstationary. For example, a transient event such as an earthquake will produce nonstationary data when measured. Nevertheless, the nonstationary character of such data is ignored or the effects assumed to be negligible. This assumption may lead to inaccurate results and incorrect interpretation of the underlying physics as explained below.

A variety of techniques have been applied to nonlinear, nonstationary physical signals. For example, many computer implemented methods apply Fourier spectral analysis to examine the energy-frequency distribution of such signals.

Although the Fourier transform that is applied by these computer implemented methods is valid under extremely general conditions, there are some crucial restrictions: the system must be linear, and the data must be strictly periodic or stationary. If these conditions are not met, then the resulting spectrum will not make sense physically.

A common technique for meeting the linearity condition is to approximate the physical phenomena with at least one linear system. Although linear approximation is an adequate solution for some applications, many physical phenomena are highly nonlinear and do not admit a reasonably accurate linear approximation.

Furthermore, imperfect probes/sensors and numerical schemes may contaminate data representative of the phenomenon. For example, the interactions of imperfect probes with a perfect linear system can make the final data nonlinear.

Many recorded physical signals are of finite duration, nonstationary, and nonlinear because they are derived from physical processes that are nonlinear either intrinsically or through interactions with imperfect probes or numerical schemes. Under these conditions, computer implemented methods which apply Fourier spectral analysis are of limited use. For lack of alternatives, however, such methods still apply Fourier spectral analysis to process such data.

In summary, the indiscriminate use of Fourier spectral analysis in these methods and the adoption of the stationarity and linearity assumptions may give inaccurate results some of which are described below.

3

First, the Fourier spectrum defines uniform harmonic components globally. Therefore, the Fourier spectrum needs many additional harmonic components to simulate nonstationary data that are nonuniform globally. As a result, energy is spread over a wide frequency range.

For example, using a delta function to represent the flash of light from a lightning bolt will give a phase-locked wide white Fourier spectrum. Here, many Fourier components are added to simulate the nonstationary nature of the data in the time domain, but their existence diverts energy to a much wider frequency domain. Constrained by the conservation of energy principle, these spurious harmonics and the wide frequency spectrum cannot faithfully represent the true energy density of the lighting in the frequency and time space.

More seriously, the Fourier representation also requires the existence of negative light intensity so that the components can cancel out one another to give the final delta function representing the lightning. Thus, the Fourier components might make mathematical sense, but they often do not make physical sense when applied.

Although no physical process can be represented exactly by a delta function, some physical data such as the near field strong earthquake energy signals are of extremely short duration. Such earthquake energy signals almost approach a delta function, and they always give artificially wide Fourier spectra.

Second, Fourier spectral analysis uses a linear superposition of trigonometric functions to represent the data. Therefore, additional harmonic components are required to simulate deformed wave profiles. Such deformations, as will be shown later, are the direct consequence of nonlinear effects. Whenever the form of the data deviates from a pure sine or cosine function, the Fourier spectrum will contain harmonics. Furthermore, both nonstationarity and nonlinearity can induce spurious harmonic components that cause unwanted energy spreading and artificial frequency smearing in the Fourier spectrum. In other words, the nonstationary, stochastic nature of biological data suffers from conventional signal processing techniques and makes the interpretation of the processed data quite difficult.

Biological Signal Analysis

According to the above background, there is a need for a more accurate signal processing technique that produces results that are more physically meaningful and readily understood. Biological signals provide another example of physical signals in which this invention is applicable. Parent application Ser. No. 08/872,586 filed Jun. 10, 1997 and now issued, U.S. Pat. No. 5,983,162, illustrates several other types of signals in which this invention is applicable. Namely, the patent provides specific examples of nonlinear, nonstationary geophysical signals which are very difficult to analyze with traditional computer implemented techniques including earthquake signals, water wave signals, tsunami signals, ocean altitude and ocean circulation signals.

Many of the aforementioned signal processing problems exist when biological signals are processed. For example, most data in the field of biology are nonstationarily stochastic. When conventional tools such as Fourier Analysis are applied to such biological data, the result often-times obscures the underlying processes. In other words, conventional Fourier analysis of biological data throws away or otherwise obscures valuable information. Thus, the complex biological phenomena producing such data cannot be readily understood and is, in any event, represented imprecisely. The interpretation of the results of such conventionally processed data may, therefore, be quite difficult. The con-

4

ventional techniques also make accurate modelling of the biological phenomena very difficult and, sometimes, impossible.

Acoustical Signal Analysis

The idea of recording and transmitting sounds and speeches depends, for example, on the variations of the density in the air, and currents in the telephone. The crucial element is on the variation, without which the sound would be a monotonic tone that would not carry any information. Speech consists of time varying acoustical signals, which are nonstationary and nonlinear. In fact, for the acoustical signal to carry any information at all, be it speech or music notes, there must be a time variation in amplitude and frequency continuously and, may be, subtly. Unfortunately, the tools we have to deal with nonstationary processes are quite limited; therefore, we are forced to make all kinds of approximations: As a result, there is a conflict between the human perception and automatic sound processes.

Some publications listed, relating to acoustical signal analysis, are incorporated by reference:

Allen, J. B., 1994: How do humans process and recognize speech? IEEE Trans. Speech and Audio Proc., 2, 567–577.

Banbrook, M, S. McLaughlin, and I. Mann, 1999: Speech characterization and synthesis by nonlinear methods. IEEE Trans. Speech and Audio Proc., 7, 1–17.

Billa, J. and A. El-Jaroudi, 1998: An analysis of the effect of basilar membrane nonlinearities on noise suppression, J. Acoust. Soc. Am., 103, 2691–2705.

Breen, A. 1992: Speech synthesis models: A Review. Electron. Commun. Eng. J., 19–31, February, 1992.

Carmona, R., W. L. Hwang, and B. Torresani 1997: Characterization of signals by the ridges of their wavelet transform, IEEE Trans. Signal Processing, 45 2586–2589.

D'Alessandro, C., V. Darsinos and B. Yegnanarayana, 1998: Effectiveness of a periodic and aperiodic decomposition method for analysis of voice sources, IEEE Trans. Speech and Audio Proc., 6, 12–23.

Furui, S. and M. M. Sondhi, 1991: Advances in Speech Signal Processing, Marcel Dekker, New York.

George, D. E. and E. Salari, 1997: Real-time pitch extraction of voiced speech, J. Network and Computer Applications, 20, 379–387.

Hoffmann, R. and C. M. Westendorf, 1997: The development of analysis methods for speech recognition. Behavioural Processes, 39, 113–125.

Huang, N. E., 2000: A New Method For Nonlinear And Nonstationary Time Series Analysis: Empirical Mode Decomposition and Hilbert Spectral Analysis. Proceedings SPIE Conference on Wavelet, Orlando, April 2000.

Huang, N. E. , Z. Shen, and S. R. Long, M. C. Wu, E. H. Shih, Q. Zheng, C. C. Tung, and H. H. Liu, 1998: The Empirical Mode Decomposition Method and the Hilbert Spectrum for Non-stationary Time Series Analysis, Proc. Roy. Soc. London, A454, 903–995.

Huang, N. E., Z. Shen, R. S. Long, 1999: A New View of Nonlinear Water Waves—The Hilbert Spectrum, Ann. Rev. Fluid Mech. 31, 417–57.

Kay, S. M. and R. Sudhaker, 1986: A zero crossing-based spectrum analyzer, IEEE Trans. Acoust. Speech, Signal Processing, 34, 96–104.

Kedem, B., 1986: Spectral analysis and discrimination by zero-crossing, Proc. IEEE, 74, 1477–1493.

US 6,862,558 B2

5

6

Kim, D. S., S. Y. Lee and R. M. Kil, 1999: Auditory processing of speech signals for robust speech regnition in real-world noisy environments, IEEE Trans. Speech and Audio Proc., 7, 55–69.

Kumar, A. and S. Mullick, 1996: Nonlinear dynamical analysis of speech, J. Acoust. Soc. Amer., 100, 615–629, 1996.

Kumaresan, R., 1998: An inverse signal approach to computing the envelope of a real valued signal, IEEE Signal Processing Letters, 5, 256–259.

Kumaresan, R. and A. Rao, 1999: Model based approach to envelope and positive instantaneous frequency estimation of signal with speech applications, J. Acoust. Soc. Am., 105, 1912–1924.

Li, T. H, and J. D. Gibson, 1996: Speech analysis and segmentation by parametric filtering, IEEE Trans. Speech and Audio Proc., 4, 203–213.

Loughlin, P. J., J. W. Pitton, and L. E. Atlas, 1994: Construction of positive time-frequency distributions, IEEE Trans. Signal Proc., 42, 2697–2705.

Loughlin, P. J., B. Tacer, 1996: On the amplitude- and frequency-modulation decomposition of signals, J. Acoust. Soc. Am., 100, 1594–1601.

Maragos, P., J. F. Kaiser, and T. F. Quatieri, 1993: Energy separation in signal modulations with application to speech analysis, IEEE Trans. Signal Processing, 41, 3024–3051.

Maragos, P. and A. Potamianos, 1999: Fractal dimensions of speech sounds: Computation and application to automatic speech recognition, J. Acoust. Soc. Am, 105, 1925–1932.

Picinbono, B., 1997: On the instantaneous amplitude and phase of signal, IEEE Trans. Signal Processing, 45, 552–560.

Pinter, Istvan, 1996: Perceptual wavelet-representation of speech signals and its application to speech enhancement. Computer Speech and language, 10, 1–22.

Pitton, J. W., L. E. Atlas, and P. J. Loughlin, Applocay-ionsof positive time-frequency distributions to speech processing. IEEE Trans. Speech and Audio Proc., 2, 554–566.

Poletti, M. A., 1997: The Homomorphic analytic signal, IEEE Trans. Signal Processing, 45, 1943–1953.

Potamianos, A. and P. Maragos, 1994: A comparison of the energy operator and Hilbert transform approach to signal and speech demodulation, Signal Processing, 37, 95–120.

Potamianos, A. and P. Maragos, 1999: Speech analysis and synthesis using an AM-FM modulation model, Speech Communication, 28, 195–209.

Quatieri, T. F., T. E. Hanna, and G. C. O'Leary, 1997: AM-FM separation using auditory-motivated filters, IEEE Trans. Speech and Audio Processing, 5, 465–480.

Rabiner, L. R. and R. W. Schafer, 1978: Digital Processing of Speech Signals, Prentice hall, Upper Saddle River, N.J. 512pp.

Rabiner, L. R. and B. H. Juang, 1993: Fundamentals of Speech Recognition, Prentice Hall, Englewoods, N.J., 507pp.

Schroeter, J. and M. M. Sondhi, 1994: Techniques for estimating vocal-tract shapes from the speech signal, IEEE Trans. Speech and Audio Proc., 2, 133–150.

Shower, E. G. and R. Biddulph, 1931: Differnttial pitch sensitivity of the ear. J. Acoust. Soc. Am, 3, 275–287.

Welling, L. and H. Ney, 1998: Formant estimation for speech recognition, IEEE Trans. Speech and Audio Proc., 6, 36–48.

Yegnanarayana, B., and R. N. J. Veldhuis, 1998: Extraction of vocal-tract system characteristics from speech signals. IEEE Trans. Speech and Audio Proc., 6, 313–327.

Yegnanarayana, B., C. d'Alessandro and V. Darsinos, 1998: An iterative althorithm for decomposition of speech signals into periodic and aperiodic components, IEEE Trans. Speech and Audio Proc., 6, 1–11.

Because our emphasis will be on speech analysis, we should first examine the principles of Human Speech Recognition (HSR) and Automatic (Machine) Speech Recognition (ASR). As summarized in the classical paper by Allen (1994), typical ASR systems start with a front end that transforms the speech signal into a feature vector. This processes is mostly through spectral analysis over a fixed period of time, within such period the speech signal is assumed to be stationary. The analysis is strictly on frequency. HSR on the other hand, processes information across frequency localized in time. Thus, the process is assumed to be nonstationary. These localized speech features are known as the formats. To extract features localized in time but across all frequencies requires time-frequency analysis. The speech signals are divided into a time-frequency continuum of formants by the cochlea of the ear, which, by its mechanical properties, is a very poor frequency discriminator. Yet, classic experiments by Shower and Biddulph (1931) has shown that the ear can detect frequency difference as small as 3 Hz near a signal of 1000 Hz. Such perceptual acuity for pitch seems to violate the 'uncertainty relation' for stationary processes, where _f _t _, with _f as the standard deviation of the frequency, and _t as the given time period. Clearly, using the traditional method based on the stationary assumption cannot explain the HSR processes. Therefore, for an accurate sound perception, we need a different paradigm of sound signal analysis, one with time-frequency analysis. This is crucial for speech recognition. It is also the cause of the long standing difficulty in Speech recognition as stated in the classic book by Rabiner and Juang (1993): "Although there is a solid basis for the linguist description of sounds and a good understanding of the associates acoustics of sound production, there is, at best, a tenuous relationship between a given linguistic sound and a repeatable, reliable, measurable set of acoustic parameters." To overcome this difficulty, we need a new signal analysis method.

To deal with the nonstationary properties of speech, various methods have been employed (see, for example, the classical book by Rabiner and Schafer (1978), and more recent developments by Furui and Sonfhi (1992), Hoffman and Westendorf (1997)), that include spectral analysis, filter-bank, zero-crossing, pattern recognition dynamical programming, linear prediction, statistic methods, and neural network. In many of these approaches, there lies a tactic assumption that the speech can be treated as locally stationary. Although great progress has been made, in speech recognition, the locally stationary assumption has rendered speech synthesis to bear a flat, wooden and artificial tone.

In addition to speech recognition, there are a wide variety of speech communications application that will depend on detailed signal analysis, such as digital transmission and storage of speech signals, speech synthesis, speaker verification and identification, and the enhancement of signal quality. The techniques used in all of these applications are not strictly limited to speeches alone, the same approach can

7

also be applied to musical scores and recordings, and even to machine condition monitoring.

So far, we have only touched on the problem of nonstationary properties of the acoustical signal. A more important part is the nonlinear properties of the acoustical signal in general and speech in particular, which has mostly been neglected up until very recently (see, for example, Kumar, 1996). As it is well known that the sound in speech is generated by three mechanisms:

(1) The vibration of the vocal chords,

(2) The friction of air through construction of the vocal tract, and

(3) The explosion of a sudden release of the air from complete closure of the vocal tract.

Of the three mechanisms, the vowels are generated by the vibration of the vocal cord with unrestricted passage of air. Such sound can be generated for indefinite length as long as the lung can supply the airflow. Therefore, the vowels are the only sounds that could be approximated locally as stationary, but only locally. Unfortunately, in our speech, vowels are not the information carrying sounds, the consonant's are. The consotant's varying of frequency is a necessity for transmitting information. For example, if we write the sentence,

Do you understand what John just said.

as,

D- y- - -nd-rst-nd wh-t J-hn -jst s- -d.

Most people can still figure out the meaning of the utterance. On the other hand, if we would write the sentence as

-o - ou u- -e- - -a- - - - a- -o- - - --- u- -ai- .

No one would be able to decipher what is its meaning at all. Consequently, in speech recognition, the precise analysis of the consonant signals is crucial.

Consonants are highly transient and the generating mechanisms are all nonlinear. Such nonlinearity has been known for a long time (See, for, example, Rabiner and Schafer, 1978), but only recently has there been any investigations (see, for example, Billa and El-Jaroudi, 1998; and Maragos and Potamianos, 1999). Although the locally stationary assumption can be used for vowels better than the consonants, we have to examine the consonants, for, as we have seen, consonants contain far more information than vowels. This creates obvious difficulties for the present approach in speech signal analysis using methods based on linear and stationary assumptions such as filter bank or spectrogram etc. Considering the generation mechanisms, we can immediately find that mechanisms involved in generating the consonants are all nonlinear. Therefore, we need a signal analysis method that is not only applicable to nonstationary but also to nonlinear processes. The Empirical Mode Decomposition (EMD) is the only method known for the task. Previously, an application of the EMD has been the Hilbert-Huang Transform (HHT), as described in previous patent applications and publications. In speech communication, we process information frequency and amplitude locally and instantaneously. The natural way is not to wait for a long string of data and then to process it spectrally. Built in with an intermittent test option, we can use EMD as a perfect tool to construct a filter bank, but the filter is in temporal space rather than in frequency space. This temporal space filtering retains the full nonlinearity

8

without resorting to the spurious harmonics. It also eliminates the distortion that could be introduced through Fourier based frequency filtering. Therefore, EMD is a natural choice. The product of EMD is actually the formant components without the difficulties of frequency resolution in any range.

Speech signal analysis is the most fundamental requirement for speech synthesis, speaker verification and identification, speech recognition, and enhancement and restoration of speech record. The basic technique can also be applied to processing music

SUMMARY OF THE INVENTION

An object of the present invention is to solve the above-mentioned problems in conventional signal analysis techniques.

Another object of the present invention is to provide further examples of physical signal processing thereby further demonstrating the broad applicability of the invention to a wide array of physical signals which include acoustical signals.

Another object is to provide a technique of distilling a physical signal to the point at which the signal can be represented with an analytic function.

To achieve these objects, the invention employs a computer implemented Empirical Mode Decomposition method which decomposes physical signals representative of a physical phenomenon into components. These components are designated as Intrinsic Mode Functions (IMFs) and are indicative of intrinsic oscillatory modes in the physical phenomenon.

Contrary to almost all the previous methods, this new computer implemented method is intuitive, direct, a posteriori, and adaptive, with the basis of the decomposition based on and derived from the physical signal. The bases so derived have no close analytic expressions, and they can only be numerically approximated in a specially programmed computer by utilizing the inventive methods disclosed herein.

More specifically, the general method of the invention includes two main components or steps to analyze the physical signal without suffering the problems associated with computer implemented Fourier analysis, namely inaccurate interpretation of the underlying physics or biology caused in part by energy spreading and frequency smearing in the Fourier spectrum.

The first step is to process the data with the Empirical Mode Decomposition (EMD) method, with which the data are decomposed into a number of Intrinsic Mode Function (IMF) components. In this way, the signal will be expanded by using a basis that is adaptively derived from the signal itself.

The second step of the general method of the present invention is to apply the Hilbert Transform to the decomposed IMF's and construct an energy-frequency-time distribution, designated as the Hilbert Spectrum, from which occurrence of physical events at corresponding times (time localities) will be preserved. There is also no close analytic form for the Hilbert Spectrum. As explained below, the invention avoids this problem by storing numerical approximations in the specially programmed computer by utilizing the inventive method.

The invention also utilizes instantaneous frequency and energy to analyze the physical phenomenon rather than the global frequency and energy utilized by computer implemented Fourier spectral analysis.

Furthermore, a computer implementing the invention, e.g., via executing a program in software, to decompose physical signals into intrinsic mode functions with EMD and generate a Hilbert spectrum is also disclosed. Because of the lack of close form analytic expression of either the basis functions and the final Hilbert spectrum; computer implementation of the inventive methods is an important part of the overall method, however other implementation of the invention may be done.

Still further, the invention may take the form of an article of manufacture. More specifically, the article of manufacture is a computer-usable medium, including a computer-readable program code embodied therein wherein the computer-readable code causes a computer to execute the inventive method.

Once the IMF's are generated, the invention can then produce a distilled or otherwise filtered version of the original physical signal. This distillation process eliminates undesired IMF's and thereby generates a filtered signal from which it is possible to perform a curve fitting process. In this way, it is possible to arrive at an analytic function which accurately represents the physically important components of the original signal.

This invention discloses a method in which Hilbert spectrum generated from IMF decomposed from a first signal may be compared with another Hilbert Spectrum to determine identification of the first signal.

Further scope of applicability of the present invention will become apparent from the detailed description given hereinafter. However, it should be understood that the detailed description and specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description. Furthermore, all the mathematic expressions are used as a short hand to express the inventive ideas clearly and are not limitative of the claimed invention.

BRIEF DESCRIPTION OF DRAWINGS

The present invention will become more fully understood from the detailed description given hereinbelow and the accompanying drawings which are given by way of illustration only, and thus are not limitative of the present invention, and wherein:

The file of this patent contains at least one drawing executed in color. Copies of this patent with color drawing(s) will be provided by the Patent and Trademark Office upon request and payment of the necessary fee.

FIG. 1(a) is a high-level flowchart describing the overall inventive method which may be implemented on the computer system shown in FIG. 2;

FIG. 1(b) is a high-level flowchart describing the Sifting Process which may be implemented on the computer system shown in FIG. 2;

FIG. 1(c) is a continuation of the high-level flowchart in FIG. 1(b) describing the Sifting Process which may be implemented on the computer system shown in FIG. 2;

FIG. 1(d) is a high-level flowchart describing EMD signal filtering and curve fitting which may be implemented on the computer system shown in FIG. 2;

FIG. 2 is a high-level block diagram of a computer system which may be programmed with the inventive with the result being a special purpose computer;

FIG. 3(a) shows wind speed data in the form of a graph plotting wind speed as a function of time for explaining the

computer implemented Empirical Mode Decomposition method of the invention;

FIG. 3(b) is a graph illustrating the upper envelope, lower envelope, envelope mean and original wind speed data which are utilized to explain the computer implemented Empirical Mode Decomposition method of the invention;

FIGS. 3(c)–(e) are graphs of the first, second and third component signals h1, h11, h12, respectively which are generated by the Sifting Process of the invention;

FIG. 3(f) is a graph of the first intrinsic mode function component which is generated by the Sifting Process of the invention;

FIG. 3(g) is a graph of data with intermittency for illustrating an optional intermittency test of the invention;

FIGS. 3(h)–(j) are graphs of the first, second, and third intrinsic mode functions when the Sifting Process is applied to the data of FIG. 3(g) without applying the intermittency test option;

FIGS. 3(k)–(m) are graphs of the first, second, and third intrinsic mode functions when the Sifting Process is applied to the data of FIG. 3(g) which applies the intermittency test option.

FIG. 4(a) is a graph of a wind speed signal which is for explaining the computer implemented Empirical Mode Decomposition method of the invention;

FIGS. 4(b)–(k) show the wind speed signal and the nine intrinsic mode functions which are extracted from the wind speed signal by the EMD method of the invention;

FIGS. 5(a)–(j) are a series of graphs illustrating the successive reconstruction of the original wind speed data from the intrinsic mode functions;

FIG. 6(a) is the Hilbert Spectrum generated by the invention from the wind speed data of FIG. 4(a);

FIG. 6(b) is the conventional Morlet Wavelet spectrum generated from the wind speed data of FIG. 4(a);

FIG. 6(c) shows the Hilbert Spectrum of FIG. 6(a) after smoothing by a 15×15 weighted Gaussian smoothing filter;

FIG. 7 is a comparison of the marginal Hilbert spectrum (solid line) and the Fourier spectrum (dotted line) which were generated from the wind speed signal of FIG. 4(a);

FIG. 8(a) is a graph illustrating the Degree of Stationarity and Degree of Statistical Stationarity which were generated from the wind speed signal of FIG. 4(a) with time averages of 10, 50, 100 and 300; FIGS. 8(b) and (c) are sections of the wind speed data that was used by the invention to produce the Degree of Stationarity shown in FIG. 8(a);

FIGS. 9(a)–(d), (f) and (g) are graphs of blood pressure data taken from the pulmonary artery of an normal, active rat which provide examples of biological data that may be processed by the invention;

FIG. 9(e) shows an envelope linking the systolic pressure extrema values for explaining the concepts of the invention;

FIG. 10(a) shows a conventional Fourier Spectrum (energy versus frequency) of the blood pressure data from FIG. 9(b) for illustrating advantages of the invention;

FIGS. 10(b)–(c) show conventional Fourier Spectrums (energy versus frequency) of the blood pressure data from FIGS. 9(c)–(d) for further illustrating advantages of the invention;

FIG. 10(d) shows a conventional Fourier Spectrum (time versus frequency) of the blood pressure data from FIG. 9(b) for illustrating advantages of the invention;

FIGS. 10(e)–(f) show conventional three-dimensional Fourier Spectrum (amplitude of spectrum as a function of

frequency in every 1-minute window on the time-frequency plane in linear scales) of the blood pressure data from FIG. 9(b) for illustrating advantages of the invention;

FIG. 10(g) is a combined graph directly comparing a conventional Fourier Spectrum and a marginal Hilbert Spectrum according to the invention calculated from the data of FIG. 9(c);

FIGS. 11(a)–(h) are graphs of the first through eighth intrinsic mode functions which are extracted from the blood pressure signal of FIG. 9(c) by the EMD method of the invention;

FIGS. 12(a)–(h) are graphs of the first through eighth intrinsic mode functions which are extracted from the blood pressure signal of FIG. 9(d) by the EMD method of the invention;

FIGS. 13(a) and (b) are reconstructions of the blood pressure signal of FIG. 9(d) based on subsets of the intrinsic mode functions;

FIG. 13(c) is another reconstructions of the blood pressure signal of FIG. 9(d) based on a different subset of the intrinsic mode functions plotted together with the original signal (dotted line) of FIG. 9(d);

FIG. 14(a) is a Hilbert Spectrum of the FIG. 9(d) blood pressure signal calculated according to the invention;

FIG. 14(b) is a Hilbert Spectrum of the FIG. 9(c) blood pressure signal calculated according to the invention;

FIGS. 15(a)–(d) are graphs of pulmonary blood pressure signals in response to step changes in oxygen concentration in the breathing gas;

FIGS. 15(e)–(h) illustrate the inventive sifting process as it is applied to the data of FIGS. 15(b);

FIGS. 16(a)–(p) are graphs of the first through sixteenth intrinsic mode functions which are extracted from the blood pressure signal of FIG. 15(a) by the EMD method of the invention;

FIGS. 17(a)–(f) are mean trends of pulmonary arterial blood pressure which are computed according to the invention;

FIGS. 18(a)–(c) are analytic functions derived by the invention and representing the indicial response of pulmonary arterial blood pressure to a step decrease in oxygen concentration from 20.9 to 10.0%;

FIGS. 19(a)–(c) are analytic functions derived by the invention and representing the indicial response of pulmonary arterial blood pressure to a step increase in oxygen concentration from 10.0 to 20.9%;

FIGS. 20(a)–(d) are oscillations about the mean trend as defined by the invention for k=1, 2, 4 and 6, respectively;

FIG. 21 shows the Hilbert energy spectrum $E_k(t)$ according to the invention which is calculated from the instantaneous amplitude spectrum of the oscillations about the mean $X_k(t)$ with k=6 (the data from FIG. 20(d));

FIGS. 22(a)–(h) are graphs of the first through sixteenth intrinsic mode functions which are extracted from the Hilbert energy spectrum $E_k(t)$ of FIG. 21 by the EMD method of the invention;

FIGS. 23(a)–(b) show the three dimensional (amplitude-frequency-time) and two-dimensional (contour of amplitude on the frequency-time plane) plot of the Hilbert Spectrum (HHT) taken from the data of FIG. 20(d);

FIG. 24 shows a conventional two dimensional Fourier Spectrum (FFT) of the pressure signal in 1-minute segments under the assumption that the process is stationary in each segment;

FIG. 25 is a graph of heart pulse interval versus time taken from a human with sleep apnea and including both a normal and abnormal (apnea) data section;

FIG. 26 is a graph of a normal heart pulse interval versus time;

FIGS. 27(a)–(h) are graphs of the first through seventh intrinsic mode functions and the residue which are extracted from the normal heart pulse interval data of FIG. 26 by the EMD method of the invention;

FIG. 28 is a blow up of FIG. 26;

FIG. 29 is a Hilbert Spectrum of the FIG. 26 normal heart pulse interval data calculated according to the invention;

FIG. 30 is a graph of an abnormal heart pulse interval versus time;

FIGS. 31(a)–(h) are graphs of the first through seventh intrinsic mode functions and the residue which are extracted from the abnormal heart pulse interval data of FIG. 29 by the EMD method of the invention;

FIG. 32 is a blow up of FIG. 30;

FIG. 33 is a Hilbert Spectrum of the FIG. 30 abnormal heart pulse interval data calculated according to the invention;

FIG. 34 is heart pulse rate data for 45 minute interval during which the patient suffered an epileptic seizure (at time index 0);

FIGS. 35(a)–(i) are graphs of the first through ninth intrinsic mode functions which are extracted from the heart pulse interval data of FIG. 34 by the EMD method of the invention;

FIGS. 36(a)–(c) are the first intrinsic mode function the original heart pulse interval data, and the third intrinsic mode function of the FIG. 34 epileptic seizure data plotted on an expanded scale, respectively;

FIG. 37 is a Hilbert Spectrum of the FIG. 34 epileptic seizure, heart pulse interval data calculated according to the invention; and

FIG. 38 is a conventional Wavelet Spectrum of the FIG. 34 epileptic seizure, heart pulse interval data for comparison with the inventive result of FIG. 37.

FIG. 39 is a block diagram describing the inventive method of speech analysis.

FIG. 40 is a block diagram describing the inventive method of speech synthesis.

FIG. 41 is a block diagram describing the inventive method of speaker identification.

FIG. 42 is a block diagram describing the inventive method of speech recognition.

FIG. 43 is a block diagram describing the inventive method of sound quality enhancement.

FIG. 44 is a block diagram describing the inventive method of machine health monitoring.

FIG. 45(a) shows acoustical signal data recorded from a male speaker, saying 'Halloo.'

FIG. 45(b) shows acoustical signal data recorded from a female speaker, saying 'Halloo.'

FIGS. 46(a)–(b) show the EMD decomposed IMF components of FIGS. 45(a)–(b).

FIGS. 47(a)–(b) show the Hilbert Spectra of FIGS. 46(a)–(b).

FIGS. 48(a)–(b) show the Fourier Spectra of FIGS. 46(a)–(b).

FIGS. 49(a)–(b) show the detailed Hilbert Spectra of FIGS. 46(a)–(b).

15

activity, process or phenomenon is sensed by an appropriate sensor in step **100**. Appropriate sensors for detecting the physical activity and generating a physical signal representative thereof are discussed in the practical examples below. As an equivalent alternative, the physical signal can be inputted in step **100**.

After sensing in step **100**, the analog signal is converted to the digital domain suitable for computer processing in the A/D conversion step **105**. Depending upon whether the input signal is analog or digital step **105** may be bypassed.

Next, an optional smoothing step **110** may be applied to the physical signal. The optional smoothing step **110** may be applied to smooth the signal with, for example, a weighted running average to remove excessive noise.

Thereafter, the Sifting Process is applied in step **120** to Sift the signal with the Empirical Mode Decomposition method and thereby extract the intrinsic mode function(s). The intrinsic mode functions can then be displayed as shown in step **130** and checked for orthogonality in step **135**.

Before continuing with the main flow in FIG. 1(*a*), the details of the Sifting Process will be explained with reference to the high level flowchart in FIGS. 2(*a*), 2(*b*) and the series of graphs showing illustrative results of the Sifting Process in FIGS. 3(*a*)–(*f*).

As shown in FIG. 1(*b*), the digitized physical signal from step **105** is first windowed by framing the end points in step **107**. Then, the Sifting Process begins at step **200** by identifying local maximum values of the digitized, framed physical signal from step **107**. FIG. 3(*a*) shows a typical physical signal **10** which, in this example, represents wind speed spanning a time interval of one second.

Before construction of the envelope in steps **210** and **230**, optional intermittency tests (**201,221**) may be introduced to alleviate the alias associated with intermittence in the data that can cause mode mixing.

Optional intermittency test **201** checks the distance between successive maxima to see if this distance between is within a pre-assigned value n times the shortest distance between waves. If no, then an intermittency exists and the method proceeds to step **202**. If yes, then there is no intermittency and the upper envelope is constructed in step **210** as further described below.

Similarly optional intermittency test **221** checks the distance between successive minima to see if this distance is within a pre-assigned value n times the shortest distance between waves. If no, then an intermittency exists and the method proceeds to step **222**. If yes, then there is no intermittency and the upper envelope is constructed in step **230** as further described below.

An example of such intermittency is given in FIG. 3(*g*), in which small scale waves appear only intermittently. By strict application of the Sifting Process, the resulting IMFs are given in FIGS. 3(*b*)–(*j*), in which two drastically different time scales are present in the first IMF component as shown in FIG. 3(*h*). This mixing of modes breaks up the main wave train by the small intermittent oscillations.

If intermittency tests (**201,222**) are employed which utilize a preassigned value of n-times the shortest distance between waves, the resulting IMFs are shown in FIGS. 3(*k*)–(*m*), in which the modes are clearly and physically separated. The effective step to eliminate the mode mixing is achieved by treating the local extrema which failed the intermittency test as local maxima and minima (steps **202** and **212**), respectively. Therefore, the upper and lower envelope will be identical as the original data reference line.

These intermittency tests (**201,221**) and the further steps (**202,222**) are optional. By selecting an artificially large n

16

value utilized in steps **201** and **221** to test for intermittency, the test will be effectively bypassed. Otherwise, the test can be bypassed at the initial selection of the program.

Then, the method constructs an upper envelope **20** of the physical signal **10** in step **210**. The upper envelope **20** is shown in FIG. 3(*b*) using a dot-dashed line. This upper envelope **20** is preferably constructed with a cubic spline that is fitted to the local maxima.

Next, the local minimum values of the physical signal **10** are identified in step **220**. To complete the envelope, a lower envelope **30** is constructed from the local minimum values in step **230**. The lower envelope **30** is also shown in FIG. 3(*b*) using a dot-dash line. Like the upper envelope **20**, the lower envelope **30** is preferably constructed with a cubic spline that is fitted to the local minima.

The upper and lower envelopes **20, 30** should encompass all the data within the physical signal **10**. From the upper and lower envelopes **20, 30**, an envelope mean **40** is the determined in step **240**. The envelope mean **40** is the mean value of the upper and lower envelopes **20, 30**. As can be seen in FIG. 3(*b*), the envelope mean **40** bisects the physical signal **10** quite well.

Then, the method generates the first component signal h_1 in step **250** by subtracting the envelope mean **40** from the physical signal **10**. This computer implemented step may also be expressed as:

$$X(t)-m_1=h_1 \qquad 1$$

Where the envelope mean **40** is m_1 and the physical signal is X(t).

FIG. 3(*c*) shows the first component signal h_1. Ideally, the first component signal h_1 should be an IMF, for the construction of h_1 described above seems to have made h_1 satisfy all the requirements of an IMF. In reality, however, a gentle hump that resides on a slope region of the data can become an extremum when the reference coordinate is changed from the original rectangular coordinate to a curvilinear coordinate. For example, the imperfection of the envelopes **20, 30** can be seen by observing the overshoots and undershoots at the 4.6 and 4.7 second points in FIG. 3(*b*).

An example of this amplification can be found in the gentle hump between the 4.5 and 4.6 second range in the data in FIG. 3(*a*). After the first round of sifting, the gentle hump becomes a local maximum at the same time location in the first component signal h_1 shown in FIG. 3(*c*). New extrema generated in this way actually recover the proper modes lost in the initial examination. Thus, the Sifting Process extracts important information from the signal which may be overlooked by traditional techniques. In fact, the Sifting Process can recover low amplitude riding waves, which may appear as gentle humps in the original signal, with repeated siftings.

Still another complication is that the envelope mean **40** may be different from the true local mean for nonlinear data. Consequently, some asymmetric wave forms can still exist no matter how many times the data are sifted. This must be accepted because, after all, the inventive method is an approximation as discussed before.

Other than these theoretical difficulties, on the practical side, serious problems of the spline fitting can occur near the ends, where the cubic spline being fit can have large swings. Left by themselves, the end swings can eventually propagate inward and corrupt the whole data span especially in the low frequency components. A numerical method has been devised to eliminate the end effects details of which will be given later. Even with these problems, the Sifting Process can still extract the essential scales from the data.

The Sifting Process serves two purposes: to eliminate riding waves and to make the wave profiles more symmetric. Toward these ends, the Sifting Process has to be repeated. Because only the first component signal h_1 has been generated so far, the decision step **260**, which tests successive component signals to see if they satisfy the definition of an IMF, is bypassed during the first iteration.

Thus, step **265** is performed which treats the component signal as the physical signal in the next iteration. The next iteration is then performed by executing steps **200–250**. In step **250**, the second component signal h_{11} is generated by subtracting the envelope mean from the physical signal (in this iteration, the first component signal h_1 is treated as the physical signal). In more formal terms:

$$h_1 - m_{11} = h_{11} \qquad\qquad 2$$

FIG. **3**(*d*) shows the second component signal h_{11}. Although the second sifting shows great improvement in the signal with respect to the first sifting, there is still a local maximum below the zero line. After a third sifting, the result (third component signal h_{12}) is shown in FIG. **3**(*d*). Now all the local maxima are positive, and all the local minima are negative, but to ensure this configuration is stable, the Sifting Process should be further repeated. In general, the Sifting Process is repeated at least 3 more times and, in general, K times to produce h_{1k}. If no more new extrema are generated, then h_{1k} is an IMF. In formal terms:

$$h_{1(k-1)} - m_{1k} = h_{1k}; \qquad\qquad 3$$

The resulting first IMF component is shown in FIG. **3**(*f*) after 9 siftings. The first IMF component of the physical signal may be designated as such in step **270** and stored in step **275** in memory **415**:

$$c_1 = h_{1k}, \qquad\qquad 4$$

As mentioned above, all these manipulations are carried out numerically in a computer **410**. There is not explicit close form analytic expression for any of the computer implemented steps.

As described above, the process is indeed like sifting of the data by the computer **410** because it separates the finest (shortest time scale) local mode from the data first based only on the characteristic time scale. The Sifting Process, however, has two effects:

a. to eliminate riding waves, and

b. to ensure the envelopes generated by maxima and minima are symmetric.

While the first condition is necessary for the instantaneous frequency to be meaningful (as discussed below), the second condition is also necessary in case the neighboring wave amplitudes have too large a disparity.

Unfortunately, the effects of the second condition, when carried to the extreme, could obliterate the physically meaningful amplitude fluctuations. Therefore, the Sifting Process should be applied with care, for carrying the process to an extreme could make the resulting IMF a pure frequency modulated signal of constant amplitude.

To guarantee that the IMF component retains enough physical sense of both amplitude and frequency modulations, a stopping criterion is employed to stop the generation of the next IMF component.

This stopping criterion is part of the computer implemented method and is shown as step **260** in FIG. **1**(*c*). Step **260** is a decision step that decides whether the stopping criteria has been satisfied. The preferred stopping criteria determines whether three successive component signals satisfy the definition of IMF. If three successive component signals all satisfy the definition of IMF, then the Sifting

Process has arrived at an IMF and should be stopped by proceeding to step **270**. If not, step **260** starts another iteration by proceeding to step **265** as described above.

Alternatively, another stopping criteria could be used that determines whether successive component signals are substantially equal. If successive component signals are substantially equal, then the Sifting Process has arrived at an IMF and should be stopped by proceeding to step **270**. If not, step **260** starts another iteration by proceeding to step **265** as described above.

Determining whether successive component signals are substantially equal in the alternative stopping criteria limits the size of the standard deviation, sd, computed from the two consecutive sifting results as follows:

$$sd = \sum_{t=0}^{T} \left[\frac{\Omega(h_{1(k-1)}(t) - h_{1k}(t))\Omega^2}{h_{1(k-1)}^2(t)} \right]. \qquad\qquad 5$$

A very rigorous and preferred value for sd is set between 0.2 and 0.3. Of course, if faster processing is desired, then a trade-off such as a less rigorous value for sd may be used.

Overall, the first IMF component c_1 should contain the finest scale or the shortest period component of the physical signal **10**.

Before extracting the next IMF component, a test should be conducted to determine if the Sifting Process should stop. The stopping criteria is shown in Step **300**. Step **300** determines whether the component signal has more than 2 extrema. If not, all of the IMF's have been extracted and the Sifting Process is stopped by proceeding to step **310**. If so, then additional IMF's may be extracted by continuing the process in step **320**.

Step **270** recognizes that an IMF component has been successfully generated by the Sifting Process by setting the component signal equal to an intrinsic mode function. More specifically, step **270** causes the computer **410** to store the component signal h1k as an intrinsic mode function in memory **415**.

Then, the first IMF is separated from the physical signal in step **290** to generate a residual signal. In particular, a residual signal is generated by subtracting the intrinsic mode function from the physical signal. In formal terms:

$$X(t) - c_1 = r_1. \qquad\qquad 6$$

Because the residue, r_1, still includes information of longer period components, it is treated as the new physical data and subjected to the same Sifting Process as described above. Step **320** performs this function by treating the residual signal as the physical signal in the next iteration. Thereafter, the next iteration is performed beginning with the execution of step **200** as described above.

The Sifting Process is analogous to a mechanical sieve, except it is implemented here in specially programmed computer and applied to any digital data numerically rather than mechanically.

The Sifting Process is repeated for all the subsequent r_j's. This iterative procedure may be expressed as:

$$r_1 - c_2 = r_2,$$
$$\diamond\diamond\diamond,$$
$$r_{n-1} - c_n = r_n. \qquad\qquad 7$$

Step **300** stops the Sifting Process by proceeding to stop step **310** if the residual signal r_n has more than 2 extrema. Otherwise, the method proceeds to step **320**.

In other words, Step **310** stops the Sifting Process if the residual signal r_n is monotonically increasing or decreasing.

US 6,862,558 B2

19

This stopping criterion is based on the fact that an IMF cannot be extracted from a monotonic function. If r_n is not monotonically increasing/decreasing, then a next iteration is performed beginning with step **320**.

Even for data with zero mean, the final residue still can be different from zero. For data with a trend, the final residue should be that trend.

In summary, the Sifting Process decomposes the physical signal X(t) into a series of intrinsic mode functions and a residue which may be expressed as:

$$X(t) = \overset{n}{\underset{i=1}{A}} c_i + r_n. \qquad 8$$

In other words, the invention extracts a series of IMFs by Sifting the physical signal with a computer implemented Empirical Mode Decomposition method. This IMF series cannot be generated or derived by any analytic method. It can only be extracted by the invention through a specially programmed computer through the Sifting Process.

EMD Filtering and Curve Fitting

FIG. 1(d) illustrates the inventive method of performing filtering and curve fitting. The method begins, as in FIG. 1(a), with sensing the physical phenomenon or inputting a signal representative thereof in step **100**. The optional smoothing **105** may then be applied to eliminate excess noise.

Then, the signal is Sifted with the Empirical Mode Decomposition Process to extract the intrinsic mode functions in step **120**. The IMFs may then be directly displayed **130**, stored **132** or transmitted **134**.

As mentioned above, fitting a curve to a signal is not always possible. The reasons include too many data points for efficient curve fitting and nonlinear data that does not admit an acceptable curve fit. The IMFs, however, extract important information from the original signal while reducing the number of data points and simplifying the resulting signal. By properly selecting only the most relevant IMFs, one can construct a filtered version of the original signal.

Step **350** performs this process of constructing a filtered signal from a subset of the IMFs. Selecting which IMFs should in the subset and which should not may be an empirical or intuitive process. Typically, the IMFs having the lowest frequency components are chosen as part of the selected subset with the higher frequency components excluded therefrom. In this way, a simplified (filtered) version of the original signal can be constructed which eliminates unnecessary and, hopefully, not physically components (IMFs).

More specifically, step **350** performs a summation of the selected IMFs. The selection of IMFs can be automated, e.g. the n lowest frequency IMFS, or manual.

Generating the filtered signal typically permits curve fitting to be successful and efficient. Once the filtered signal is generated, conventional curve fitting techniques (step **360**) such as a least squares estimation process can be utilized.

The fitted curve may then be outputted **362**, displayed **364**, stored **366** or transmitted **368**.

The filtered signal may also be subjected to further processing in step **370**. Namely, the oscillatory energy about the mean is calculated in step **370**. This calculation is essentially the same as the instantaneous energy density calculation explained below, with the calculation being

20

performed for a subset (k) of the IMFs. Further details and an example are discussed below. The result of step **370** is a signal that may be subjected to the Sifting Process to extract a set of IMFs in step **120**. Although FIG. 1(d) illustrates an infinite loop (steps **120**, **350** and **370**), at some point it will not be appropriate or possible extract further IMFs. Thus, this loop is not really infinite and is typically performed one or two times.

Computer for Implementing Inventive Method

A computer suitable for programming with the inventive method is diagrammatically shown in the block diagram of FIG. **2**. The computer **410** is preferably part of a computer system **400**. To allow human interaction with the computer **410**, the computer system includes a keyboard **430** and mouse **435**. The computer programmed with the inventive method is analogous to a mechanical sieve: it separates digital data into series of IMF's according to their time scales in a manner analogous to a mechanical sieve which separates aggregated sand particles according to their physical size.

Because the invention is applied to analyze physical signals, the computer system **400** also includes an input device **440**, sensor **442** and/or probe **444** which are used to sample a physical phenomenon and generate physical signal representative thereof. More particular examples of such inputs (**440**, **442** and **444**) are described in relation to FIGS. 21–25.

To output the results of the computer implemented method, the computer system **400** also includes a display **450** such as a cathode ray tube or flat panel display, printer **460** and output device **470**. Each of these outputs (**450**, **460**, **470**) should have the capability to generate or otherwise handle color outputs because, for example, the Hilbert Spectrum may be in color.

The generalized output device **470** may also include a network connection to connect the computer **400** to a local or wide area network. In this way, the physical signal may be inputted from the network. Furthermore, all outputs can be sent to another location via such a network connection.

Furthermore, the computer system **400** also includes a mass storage device **420**. The mass storage device **420** may be a hard disk, floppy disc, optical disc, etc. The mass storage device **420** may be used to store a computer program which performs the invention when loaded into the computer **410**. As an alternative, the input device **440** may be a network connection or off-line storage which supplies the computer program to the computer **410**.

More particularly, the computer program embodiment of the invention may be loaded from the mass storage device **420** into the internal memory **415** of the computer **410**. The result is that the general purpose computer **410** is transformed into a special purpose machine that implements the invention.

Even more particularly, each step of inventive method will transform at least a portion of the general purpose computer **410** into a special purpose computer module implementing that step. For example, when the sifting step **120** is implemented on the computer **410**, the result is a computer implemented sifting apparatus (sifter) that performs the sifting functions of sifting step **120**.

Other embodiments of the invention include firmware embodiments and hardware embodiments wherein the inventive method is programmed into firmware (such as EPROM, PROM or PLA) or wholly constructed with hardware components. Constructing such firmware and hardware

embodiments of the invention would be a routine matter to one of ordinary skill using known techniques.

Article of Manufacture

Still further, the invention disclosed herein may take the form of an article of manufacture. More specifically, the article of manufacture is a computer-usable medium, including a computer-readable program code embodied therein wherein the computer-readable code causes computer **410** to execute the inventive method.

A computer diskette such as disc **480** in FIG. **2** is an example of such a computer-usable medium. When the disc **480** is loaded into the mass storage device **480**, the computer-readable program code stored therein is transferred into the computer **410**. In this way, the computer **410** may be instructed to perform the inventive methods disclosed herein.

Illustration of Sifting Process

To further illustrate the Sifting Process, consider FIG. **4**(*a*) which shows a physical signal representing wind data collected in a laboratory wind-wave tunnel with a high frequency response Pitot tube located 10 cm above the mean water level. The wind speed was recorded under the condition of an initial onset of water waves from a calm surface. Clearly, the physical signal is quite complicated with many local extrema but no zero-crossings such that the time series represents all positive numbers.

Although the mean can be treated as a zero reference, defining it is difficult, for the whole process is transient. This example illustrates the advantage of adopting the successive extrema for defining the time scale and the difficulties of dealing with nonstationary data. In fact, a physically meaningful mean for such data is impossible to define using standard methods. The EMD eliminates this difficulty.

FIG. **4**(*b*) shows the wind speed signal of FIG. **4**(*a*) on a different scale for comparison purposes. FIGS. **4**(*c*)–(*k*) show all the IMFs obtained from repeatedly sifting the wind speed signal in FIG. **4**(*b*). The efficiency of the invention is also apparent: the Sifting Process produces a total of 9 intrinsic mode function components while the Fourier transform needs components which number as many as half of the total number of points in a given window to represent the wind data with the same accuracy.

The separation of the wind speed data into locally non-overlapping time scale components is clear from FIGS. **4**(*c*)–(*k*). In some components, such as c_1 and c_3, the signals are intermittent, then the neighboring components might include oscillations of the same scale, but signals of the same time scale would never occur at the same locations in two different IMF components.

The components of the EMD are usually physical, for the characteristic scales are physically meaningful.

To confirm the validity and completeness of the Sifting Process, the wind speed signal can be reconstructed from the IMF components. FIGS. **5**(*a*)–(*i*) show this reconstruction process starting from the longest period IMF to the shortest period IMF in sequence. For example, FIG. **5**(*a*) shows the wind speed signal and the longest period component, c_9, which is actually the residue trend, not an IMF.

By itself, the fitting of the trend is quite impressive, and it is very physical: the gradual decrease of the mean wind speed indicates the lack of drag from the calm surface initially and the increasing of drag after the generation of wind waves. As the mean wind speed deceases, the ampli-

tude of the fluctuation increases, another indication of wind-wave interactions.

By adding the next longest period component, c_8, the trend of the sum, c_9+c_8, takes a remarkable turn, and the fitting to the wind speed signal looks greatly improved as shown in FIG. **5**(*b*). Successively adding more components with increasing frequency results in the series of FIGS. **5**(*c*)–(*i*). The gradual change from the monotonic trend to the final reconstruction is illustrative by itself. By the time the sum of IMF components reaches c_3 in FIG. **5**(*g*), essentially all the energy contained in the wind speed signal is recovered. The components with the highest frequencies add little more energy, but they make the data look more complicated. In fact, the highest frequency component is probably not physical, for the digitizing rate of the Pitot tube is too slow to capture the high frequency variations. As a result, the data are jagged artificially by the digitizing steps at this frequency. The difference between the original data and the re-constituted set from the IMF's is given in FIG. **5**(*j*). The magnitude of the difference is 10^{-15}, which is the limit of the computer **410**.

The Hilbert Spectrum

Having obtained the Intrinsic Mode Function components (through either the local extrema or curvature extrema Sifting Processes), the next step in the computer implemented method is to apply the Hilbert Transform to each component and generate the Hilbert Spectrum as shown in step **140** in FIG. **1**(*a*). A brief tutorial on the Hilbert transform with emphasis on its physical interpretation can be found in Bendat and Piersol, 1986, *"Random Data: Analysis and Measurement Procedures"*, 2nd Ed., John Wiley & Sons, New York, N.Y. Essentially, the Hilbert transform is the convolution of X(t) with 1/t. This convolution emphasizes the local properties of X(t). In more formal terms, given a time series X(t), the Hilbert Transform Y(t) can be expressed as

$$Y(t) = \frac{1}{p} P \int_{-\infty}^{\infty} \frac{X(t\cent)}{t - t\cent} \, dt\cent \qquad 9$$

where P indicates the Cauchy principal value.

With this definition, X(t) and Y(t) form a complex conjugate pair. This complex conjugate pair Z(t) may be expressed as:

$$Z(t)=X(t)+iY(t)=a(t)e^{iq(t)}, \qquad 10$$

in which

$$a(t) = [X^2(t) + Y^2(t)]^{\frac{1}{2}}, \qquad 11$$

$$q(t) = \arctan\frac{X(t)}{Y(t)}. \qquad 12$$

After performing the Hilbert transform to each IMF component, we can express the time series data X(t) in the following form:

$$X(t) = \overset{n}{\underset{j=1}{A}} \, a_j(t) e^{i \int w_j(t) dt}. \qquad 13$$

In Equation (13), the residue, r_n, is purposefully omitted, for it is either a monotonic function, or a constant. Although the Hilbert transform can treat the monotonic trend as part

23

of a longer oscillation, the energy involved in the residual trend could be overpowering. In consideration of the uncertainty of the longer trend, and in the interest of information contained in the other low energy and higher frequency components, the final non-IMF component should be left out. It, however, could be included, if physical considerations justify its inclusion.

Note that Equation (13) gives both amplitude and frequency of each component as functions of time. It should be pointed out that no analytical method can generate the expression in Equation (13). Instead, all the components may be extracted only by a specially programmed computer applying the inventive Sifting Process and the Hilbert transform. The variable amplitude and frequency have not only greatly improved the efficiency of the expansion, but also enabled the expansion to accommodate nonstationary data. With IMF expansion, the amplitude and the frequency modulations are also clearly separated.

Equation (13) also enables the computer implemented method to represent the amplitude and frequency as functions of time in a three-dimensional plot, in which the amplitude can be contoured on the frequency-time plane. This frequency-time distribution of the amplitude is designated as the Hilbert Amplitude Spectrum, $H(\_, t)$, or simply Hilbert Spectrum. Thus we have:

$$H(w, t) = \overset{n}{\underset{j=1}{A}} a_j(t)e^{i \int w_j(t)dt}. \qquad 14$$

In which $H(\_, t)$ is the Hilbert spectrum of the frequency $(\_)$ and time (t) and $a_j(t)$ is the j-th component of the IMF.

In the presentation, the amplitude (with or without smoothing) can be expressed in color maps, black-grey maps, or contour maps. Color maps, however, greatly increase the operator's ability to fully analyze the spectrum. In some cases, a color map will permit the operator to discern relationships and trends that would not be apparent in black-grey maps thereby making a color display a necessary component in some cases.

If amplitude squared is more desirable to represent energy density, then the squared values of amplitude can be substituted to produce a Hilbert Energy Spectrum just as well.

Various forms of Hilbert spectra presentations can be generated by the computer in the display step **190**: both color coded maps and contour maps may be employed to present the Hilbert spectra with or without smoothing. The Hilbert Spectrum in the color map format for the wind data is shown in FIG. **6**(*a*). The Hilbert spectrum in FIG. **6**(*a*) gives a very different appearance when compared with the corresponding Wavelet spectrum shown in FIG. **6**(*b*). While the Hilbert Spectrum in FIG. **6**(*a*) appears only in the skeleton form with emphasis on the frequency variations of each IMF, the Wavelet analysis result gives a smoothed energy contour map with a rich distribution of higher harmonics.

If a more continuous form of the Hilbert Spectrum is preferred, a smoothing method can be optionally applied in step **155**. The first type of a smoothing method which may be used in the invention is a weighted spatial filter which averages over a range of cells. For example, a 15 by 15 weighted Gaussian filter may be employed in step **155** as is known in the art to smooth this data. FIG. **6**(*c*) shows the result of applying the 15 by 15 weighted Gaussian filter.

Although smoothing step **155** degrades both frequency and time resolutions, the energy density and its trends of evolution as functions of frequency and time are easier to identify. In general, if more quantitative results are desired, the original skeleton presentation is better. If more qualita-

24

tive results are desired, the smoothed presentation is better. As a guide, the first look of the data is better in the smoothed format.

The alternative of the spatial smoothing is to select a lower frequency resolution and leave the time axis undisturbed. The advantages of this approach are the preservation of events' locations and a more continuous frequency variation. Furthermore, a lower frequency resolution saves computation time for the computer implemented method.

To optimize such computation time, the optimal frequency resolution in the Hilbert spectrum can be computed as follows. Let the total data length be T, and the digitizing rate of the sensor be $\_t$. Then, the lowest frequency that can be extracted from the data is 1/T Hz, which is also the limit of frequency resolution for the data. The highest frequency that can be extracted from the data is $1/(n \_t)$ Hz, in which n represents the minimum number of $\_t$ needed to define the frequency accurately.

Because the Hilbert transform defines instantaneous frequency by differentiation, more data points are needed to define an oscillation. The absolute minimum number of data points is five for a whole sine wave. Although a whole sine wave is not needed to define its frequency, many points within any part of the wave are needed to find a stable derivative. Therefore, the maximum number of the frequency cells, N, of the Hilbert spectrum should be

$$N = \frac{\frac{1}{nDt}}{\frac{1}{T}} = \frac{T}{nDt}. \qquad 15$$

In order to make the derivative stable, the data is averaged over three adjacent cell values for the final presentation.

To illustrate, consider the wind data of FIG. **4**(*a*) which was digitized at a rate of 0.01 seconds and has a total length of 30 seconds. Therefore, the highest frequency that can be extracted is 25 Hz. The total cell size could be 600, but they have been averaged to 200 in FIG. **6**(*a*).

Marginal Spectrum

The marginal spectrum offers a measure of total amplitude (or energy) contribution from each frequency value. In other words, the marginal spectrum represents the cumulated amplitude over the entire data span.

As shown in FIG. **1**(*a*), the marginal spectrum is calculated by the computer implemented method in step **145** after applying the Hilbert Transform in step **140**. The marginal spectrum is the Hilbert Spectrum integrated through all time. In this simplification, the time coordinate is lost as in the Fourier spectral analysis, which leaves a summary of the frequency content of the event. Therefore, this presentation should only be used when the phenomena being analyzed is stationary. Formally, the marginal spectrum $h(\_)$ is defined as:

$$h(w) = \int_0^T H(w, t)dt. \qquad 16$$

Because there is no analytic expression for $H(\_,t)$, the integration can only be performed in a computer as a sum.

An example of a marginal spectrum is shown in FIG. **7**. More particularly, FIG. **7** shows the corresponding marginal spectrum of the Hilbert Spectrum given in FIG. **6**(*a*).

The frequency in either $H(\_, t)$ or $h(\_)$ has a totally different meaning from results generated by applying Fou-

25

rier spectral analysis. In the Fourier representation, the existence of energy at a frequency, _, means a component of a sine or a cosine wave persisted through the time span of the data.

In contrast, the existence of energy at the frequency, _, means only that, in the whole time span of the data, there is a higher likelihood for such a wave to have appeared locally. In fact, the Hilbert Spectrum is a weighted, non-normalized, joint amplitude-frequency-time distribution. The weight assigned to each time-frequency cell is the local amplitude. Consequently, the frequency in the marginal spectrum indicates only the likelihood of an oscillation with such a frequency exists. The exact occurrence time of that oscillation is given in the full Hilbert spectrum.

To illustrate this difference, the corresponding Fourier Spectrum of the wind speed signal is also given in FIG. 7 using a dotted line. As can be observed, there is little similarity between the Fourier spectrum and the marginal spectrum. While the Fourier spectrum is dominated by the DC term because of the non-zero mean wind speed, the marginal spectrum gives a nearly continuous distribution of energy. The Fourier spectrum is meaningless physically, because the data is not stationary. In contrast, the marginal spectrum provides a physically meaningful interpretation of the wind speed signal.

Instantaneous Frequency

There are two types of frequency modulation: the inter-wave and the intra-wave modulations. The first type is familiar because the frequency of the oscillation gradually changes as the waves disperse. Technically, in dispersive waves, the frequency is also changing within one wave, but that is generally not emphasized either for convenience, or for lack of a more precise frequency definition. The second type is less familiar, but it is also a common phenomenon: if the frequency changes from time to time within a wave its profile can no longer be a simple sine or cosine function. Therefore, any wave profile deformation from the simple sinusoidal form implies intra-wave frequency modulation.

In the past, such phenomena were treated as harmonic distortions. Nevertheless, such deformations should be viewed as intra-wave frequency modulation because the intra-wave frequency modulation is a more physically meaningful term.

In order to understand these frequency modulations, the invention applies a unique definition of instantaneous frequency. This definition stems from the EMD method and requires the signal to be reduced into IMF components. After extracting the IMF components, an instantaneous frequency value can be assigned to each IMF component. Consequently, for complicated data in which more than one IMF may be extracted by EMD, there can be more than one instantaneous frequency at a time locally.

With the Hilbert Transform, a unique definition of instantaneous frequency may be applied by the computer implemented method as illustrated by step **160**. Step **160** calculates the instantaneous frequency which is formally defined as follows:

$$w(t) = \frac{dq(t)}{dt}. \qquad 17$$

By calculating instantaneous frequency, step **160** of the invention permits the frequency value to be defined for every point with the value changing from point to point.

26

The validity and the implications of the instantaneous frequency for nonlinear signals may be analyzed by examining an idealized Stokes wave in deep water. The wave profile of such a Stokeian wave is modeled by

$$X(t)=\cos(wt+e \sin wt) \qquad 18$$

Therefore, it is a intra-wave frequency modulated signal. Approximately, equation (18) can be shown to be:

$$X(t)=(1+e/)\cos wt+e \cos 2wt \qquad 19$$

The wave profile is also shown in FIG. 9(*a*). Because the intra-wave frequency can only be approximated by harmonics in Fourier analysis, we can still have the same profile, but not the same frequency content. The wave form shows sharpened crests and rounded off troughs, which make the crests and troughs asymmetric with respect to the mean surface.

Processed with computer implemented EMD, this data yields only one IMF component as shown in FIG. 9(*b*), with a constant offset component (not shown). Although this wave has only one characteristic scale or IMF, the Wavelet analysis result shown in FIG. 9(*c*). FIG. 9(*c*) has many harmonics with two visible bands of energy corresponding to the highest order of approximations of the possible harmonics.

In contrast, the IMF data can be processed by the inventive method to give the Hilbert Spectrum shown in FIG. 9(*d*). The Hilbert Spectrum has only one frequency band centered around 0.03 Hz, the fundamental frequency of the wave train, but there is an intra-wave frequency modulation with a magnitude range of 0.02 to 0.04 Hz as the model truly represents. This intra-wave frequency modulation has been totally ignored in the past, for the traditional definition of frequency is based on the reciprocal of periodicity and Fourier Analysis.

Instantaneous Energy Density

Furthermore, the computer implemented method may also calculate the instantaneous energy density in step **150**. The instantaneous energy density, IE, is defined as:

$$IE(t)=Ú_W H^2(w, t)dw \qquad 20$$

Still further, this instantaneous energy density depends on time. Therefore, the IE can be used to check energy fluctuations.

Stationarity

To quantitatively measure the stationarity of a physical signal, the invention utilizes step **165** to calculate various stationarity measurements. Before introducing the preferred stationarity measurements, a brief review of conventional stationarity measurements is presented.

The classic definitions of stationarity are dichotomous: a process is either stationary or nonstationary. This crude definition is only qualitative in nature. Such definitions are both overly stringent and useless at the same time: few data sets can satisfy the rigor of these definitions; consequently, no one even bothers using them to check stationarity of the signal. As a result, data as nonstationary as earthquake and seismological signals are routinely treated as stationary (see, for example, Hu, et al., 1996., *Earthquake Engineering.* Chapman & Hall, London).

Sometimes, for some obviously nonstationary data, two less stringent definitions are invoked: piece-wise stationary and asymptotically stationary. These definitions are still dichotomous.

To quantify the statistical processes further, an index is needed to give a quantitative measure of how far the process deviates from stationarity. A prerequisite for such a definition is a method to present the data in the frequency-time space.

With the energy-frequency-time distribution (Hilbert Spectrum) described above, stationarity of the physical signal may be quantitatively determined. Therefore, the invention introduces an index of stationarity as follows and calculates a Degree of Stationarity in step **165**.

The first step in defining the Degree of Stationarity, DS(_), is to find the mean marginal spectrum, n(_), as

$$n(w) = \frac{1}{T}h(w) \qquad 21$$

Then, the Degree of Stationarity may be defined as:

$$DS(w) = \frac{1}{T}\int_0^{\acute{U}^T}\left(1 - \frac{H(w,t)}{n(w)}\right)^2 dt, \qquad 22$$

Again, the value of DS(_) can be determined by the computer. Therefore, the specialized computer **410** according to the invention can be treated as a stationarity meter.

For a stationary process, the Hilbert spectrum cannot be a function of time. Then, the Hilbert Spectrum will consist of only horizontal contour lines and DS(_) will then be identically zero. Only under this condition will the marginal spectrum be identical to the Fourier spectrum, then the Fourier spectrum will also make physical sense. On the other hand, if the Hilbert spectrum depends on time, the index will not be zero, then the Fourier spectrum will cease to make physical sense.

In general, the higher the index value, the more nonstationary is the process. The DS for the wind data is shown in FIG. **8**(*a*). As the index shows, the data are highly nonstationary especially for the high frequency components.

Eq. (22) defines the stationarity as a function of frequency. This is necessary because certain frequency components can be nonstationary while other components remain stationary. An example is sporadic riding wind-generated waves on otherwise uniform swell: the low frequency swell component is stationary, while the high frequency wind waves are intermittent, and hence nonstationary.

The degree of stationarity can also be a function of time implicitly, for the definition depends on the time length of integration in Eq. (22). Therefore, a process can be piece-wise stationary. On the other hand, for a singular outburst in an otherwise stationary signal, the process can be regarded as almost stationary if a long time integral is performed, but nonstationary the integral only encompasses the immediate neighborhood of the outburst.

Stationarity can be a complicated property of the process: for any signal shorter than a typical long wave period, the process may look transient. Yet as the signal length gets longer, the process can have many longer wave periods and becomes stationary. On the other hand, the signal can be locally stationary while in a long time sense nonstationary. An index is therefore not only useful but also necessary to quantify the process and give a measure of the stationarity.

The invention also calculates a Degree of Statistic Stationarity in step **165**. The degree of stationarity defined in Eq. (22) can be modified slightly to include statistically

stationary signals, for which the Degree of Statistic Stationarity, DSS(_,_T), is defined as

$$DS(w) = \frac{1}{T}\int_0^{\acute{U}^T}\left(1 - \frac{\overline{H(w,t)}}{n(w)}\right)^2 dt, \qquad 23$$

where the over-line indicates averaging over a definite but to shorter time span, _T, than the overall time duration of the data, T. For periodic motions, the _T can be the period. Such a time scale, however, is hard to define precisely for high dimensional, nonstationary dynamic systems.

Even with this difficulty, the definition for DSS could be more useful in characterizing random variables from natural phenomena. Furthermore, DSS will depend on both frequency and the averaging time span. For the wind data taken as an example, the DSS is given in FIG. **8**(*a*) with _T=10, 50, 100, and 300 time steps respectively. The results show that while the high frequency components are nonstationary, they can still be statistically stationary. Two frequency bands at 7 and 17 Hz are highly nonstationary as the DSS averaged at 100 time steps shown in FIG. **8**(*b*). These components are intermittent as can be seen in the IMF components and the marginal spectrum. A section of the original wind data is also plotted in FIG. **8**(*c*) to demonstrate that there are indeed prominent 7 and 17 Hz time scale oscillations.

Display of Selected Results

The invention displays various results of the above-described computer implemented method in step **190**. These displays are extremely useful in analyzing the underlying physics of the physical phenomenon being studied as described above. Furthermore, particular examples of these displays and the increased understanding of the underlying physics which these displays permit are discussed in the following section.

For example, the invention generates various Hilbert spectra displays in the display step **190**. As mentioned above, both color coded maps and contour maps may be employed to display the Hilbert spectra in display step **190**. In addition, the color coded maps convey information to the operator in a uniquely accessible way permitting a more thorough and deeper understanding of the physical phenomenon and may be considered as necessary to analyze some physical phenomena.

The displays generated by the invention in display step **190** are central to the invention because they allow an operator to analyze the underlying physics of the physical phenomenon being studied.

The display step **190** outputs displays to display **450**. As mentioned above, display **450** includes devices such as a cathode ray tube and a flat panel display. As an alternative, display step **290** may generate a hard copy output by utilizing printer **460** or may send the generated display to output device **470**.

Bilogical Signal Processing

The above description provides a foundation of signal analysis methodologies which has broad applicability to a wide variety of signal types. Once such applicable signal type is biological signals which are signals measured or otherwise representative of a particular biological phenomenon.

For example, any given physiological process has a number of quantities that can be measured to produce a signal or otherwise be represented by a signal. The invention

can be applied to such signals in an effort to better understand the underlying biological phenomenon. For example, the relationships between different biological variables can be studied with a precision and depth of understanding not possible before.

The invention can also be used to arrive at an analytical function representative of the biological phenomenon. By using the inventive filtering methods, the biological signal can be simplified to eliminate components not particularly relevant to the analysis and to thereby make it possible to represent the signal with a function.

Blood pressure data, such as the data shown in FIGS. $9(a)$–(g), are examples of the invention's applicability to analyzing and characterizing biological signals. This blood pressure data and the analysis that follows focusses on changes in blood pressure over the course of day.

To collect this data, a pressure probe was implanted into the pulmonary artery of a rat. Measurements were taken with the rat under controlled conditions (e.g. rat breathing normal atmosphere at sea level, quiet room, lit for 12 of the 24 hour time span, etc.). In the experiment, a Statham P23ID transducer was utilized, data collected by the computer **400**, and also recorded on a chart recorder. A standard A/D converter was utilized to converted the transduced analog signal to digital data for storage and processing by the computer **400**. It is to be understood that this particular set-up is exemplary in nature and illustrates one of the various ways in which the biological data can be obtained for processing by the invention.

FIG. **9**(a) shows the results of measuring the rat's blood pressure for a 24-hour period with the pressure waveform being sampled at a rate of 100 points/second. A one hour strip of this data is shown in FIG. **9**(b). Two separate 10-second strips of this data are shown in FIGS. **9**(c)–(d) with **9**(c) exhibiting less regularity than **9**(d). FIGS. **9**(e)–(f) are the systolic peaks and diastolic troughs for the one hour record shown in FIG. **9**(b), obtained by connecting the successive peaks and successive valleys, respectively.

In order to compare the invention against conventional methodologies, FIG. **10**(a) is presented which shows the conventional Fourier Spectrum for the one-hour data of FIG. **9**(b). FIG. **10**(a) shows this Fourier Spectrum from which one can easily identify spectral peaks at 1.5, 6.5 and 13 Hz. The 6.5 Hz peak represents the heartbeats which necessarily affect blood pressure. The 13 Hz peak is the harmonic of the 6.5 Hz heartbeat peak. The 1.5 Hz peak is probably related to respiration.

FIGS. **10**(b)–(c) show conventional Fourier spectra for the two 10-second sections of data (FIGS. **9**(c)–(d)). Fourier analysis for the one-hour data of FIG. **9**(b) is shown in FIG. **10**(d) in which the locations of the highest peaks of the Fourier spectra in every 1-minute window are plotted on the time-frequency plane. A perspective view of the windowed 1-hour and 10-sec Fourier results Fourier results are given in FIGS. **10**(e) and **10**(f), respectively and show the variation of the amplitudes of the signals.

FIG. **10**(g) directly compares the inventive results against conventional techniques. In FIG. **10**(g), the Fourier (dotted line) spectra and the Marginal Hilbert (solid line) spectra of the blood pressure data from FIG. **9**(c) are plotted together. As can be seen from this comparison, the Fourier spectra devotes a great deal of energy to the higher-harmonic components.

As explained above, the invention adopts the spacing of extrema values to analyze the data. Specifically, the sifting process is applied to the data to extract a set of intrinsic mode functions.

FIGS. **11**(a)–(h) show the results of applying the empirical mode decomposition method of the invention to the blood pressure data of FIG. **9**(c). As shown therein, the invention produces eight IMFs from this blood pressure data. The residual IMF R_8 is constant.

FIGS. **12**(a)–(h) show the results of applying the Empirical Mode Decomposition method of the invention to the blood pressure data of FIG. **9**(d). As shown therein, the invention produces eight IMFs $(C_1$–$C_8)$ from this blood pressure data. The residual IMF C_8 (a.k.a R_8) is constant.

In FIGS. **12**(a)–(h), the IMF components have very different amplitudes. The IMFs having the most energy are C_2, C_3 and C_4. The amplitude and periodicity of these three main components closely approximate their respective, constant levels.

By recognizing the main or most significant TMFs, one can reconstruct a filtered version of the signal that eliminates undesired components while continuing to faithfully represent the original signal. For example, FIG. **13**(b) shows a combination of IMFs C_2, C_3 and C_4. The sum of these components faithfully represents the original signal as can be seen by comparing the IMF sum in FIG. **13**(b) with the original signal in FIG. **9**(c).

Of course, other filters can be constructed with the IMFs. One such example is FIG. **13**(a) which shows a summation of IMFs C_1, C_2 and C_3 that were calculated from the blood pressure signal of FIG. **9**(c).

Another filtering example is shown in FIG. **13**(c) which is a sum of IMFs C_4, C_5, C_6, C_7 and C_8 shown in solid line. These components are the lower-frequency components and their sum recovers the slow variation of the pressure signal. The dotted line in FIG. **13**(c) shows the original pressure signal.

FIGS. **14**(a)–(b) show the Hilbert spectrums of the IMF components derived from the FIGS. **9**(d) and **9**(c) blood pressure signals, respectively. Specifically, The Hilbert spectrum of FIG. **14**(a) shows that the most prominent energy bands are centered at 6.5, 3.0 and 1.5 Hz. As shown, there are intra-wave frequency modulations in this spectrum which are indications of nonlinear dynamics. The wide fluctuations of frequency values in the FIG. **14**(b) Hilbert spectrum make any visual mean estimation exceedingly difficult.

By further applying the inventive methodologies, a marginal Hilbert Spectrum can be calculated by integrating the Hilbert Spectrum. An example is shown in FIG. **10**(g) which uses a solid line to plot the marginal Hilbert spectrum of the FIG. **9**(c) blood pressure signal. Prominent spectral peaks occur at 7 and 1.3 Hz.

FIG. **10**(g) also uses a dotted line to plot the corresponding Fourier Spectrum. By comparing these spectra, one can see that the mean peaks line up. However, the Hilbert spectrum clearly depicts the fluctuation of energy over frequencies without allowing the frequency of oscillation to be variable in the whole time window.

One basic difference between the conventional spectra of FIG. **10**(d) and the inventively derived spectra of FIG. **14**(a) lies in the stationarity hypothesis. In FIG. **10**(d), the Fourier spectrum assumes stationary oscillation. In FIG. **14**(a), the Hilbert spectra, does not make this assumption and is valid for nonstationary oscillations.

By using the invention, one can more readily and accurately recognize the various features of blood pressure signals. Moreover, the invention offers a more comprehensive view of the blood pressure fluctuation than the classical Fourier analysis.

The invention may also be used to study the effect of one variable on another variable. Specifically, by changing one variable of a system, measuring the effect on another variable, and then applying the invention one can arrive at a much deeper understanding of underlying system. The invention also permits the modelling or representation of data with an analytic function that was not possible with conventional techniques.

Biological systems are examples of systems which can be better understood and modelled by applying the invention. Biological systems are often studied by changing one variable and recording the changes of other variables The resulting data is often oscillatory, stochastic and non-stationary. Data having these properties are particularly susceptible to processing by the invention.

A particular example elaborated upon here is the effect of changing the breathing gas concentration of oxygen on the blood pressure of an animal. Analyzing such stochastic data to obtain crystal clear results describing the effects of oxygen concentration changes on blood pressure has been impossible with conventional techniques such as Fourier analysis.

Concrete data illustrating this example is shown in FIGS. 15(a)–(d). Specifically, the pulmonary blood pressure in the arterial trunk of two rats breathing normal air at sea level is shown in the upper trace of FIGS. 15(a)–(b) with the lower trace showing the oxygen concentration.

These same rats where then subjected to step changes in oxygen concentration: FIG. 15(c) shows a step increase and 15(d) shows a step decrease in oxygen concentration and the effect thereof on blood pressure. From this raw data, one can see certain trends. However, there is no definitive way to handle this overwhelmingly complex data to reveal the underlying physiological variations. Fourier analysis simply does not work for such nonstationary signals. By applying the invention, however, which works particularly well with nonstationary signals, one can arrive at a much deeper understanding of the underlying physiology.

As described above in detail, the invention applies a unique Sifting Process. FIGS. 15(e)–(h) illustrate this Sifting Process as it is applied to the blood pressure signal of FIG. 15(b). Specifically, FIG. 15(e) illustrates generating the envelopes (dot-dash line) by connecting successive extremas (peaks and troughs) of the signal (solid line) with cubic splines. The mean of these two envelopes is shown in FIG. 15(e) with a thick solid line and is denoted by the symbol $m_1(t)$.

The difference, $X(t)-m_1(t)$ is designated $h_1(t)$ as is shown in FIG. 15(f). From this figure, it can be seen that $h_1(t)$ has a few negative local maxima and positive local minima, suggesting that further Sifting needs to be performed before the first intrinsic mode is generated by the invention and that the quantity $h_1(t)$ is not yet a true representation of an "oscillation about the mean."

To improve this situation, two definitive requirements are imposed for a function that represents the "oscillations about the mean": (i) in the whole data set, the number of extrema and the number of zero-crossings must either be equal or differ at most by one, and (ii) at any time, the mean value of the envelope of the local maxima and the envelope of he local minima must be zero. As oscillatory function that is processed to satisfy these definitions is called an intrinsic mode function (IMF) as more fully described above.

The function $h_1(t)$ shown in FIG. 15(f) does not satisfy this definition or the requirements of an IMF. Thus, further Sifting is performed by treating $h_1(t)$ as the new data set and

repeating the Sifting Process including determining the new upper and lower envelopes, computing their new mean $m_{11}(t)$.

The difference $h_1(t)-m_{11}(t)$ is designated at $h_{11}(t)$ which is treated as the new data set. The process is repeated a number of times (FIGS. 16(g)–(h)) until it converges. The convergent result, FIG. 16(g), is designated by $C_1(t)$ and is the first intrinsic mode function, which has a zero local mean.

The difference between $X(t)$ and $C_1(t)$ is a function of time representing a "mean tread" after the first round of Sifting. It is designated as the "first residue" $R_1(t)$ $X(t)-C_1(t)=R_1(t)$

The first residue, $R_1(t)$, is again oscillatory and can be analyzed as new data by another round of Sifting in the Empirical Mode Decomposition Process, yielding the second intrinsic mode function $C_2(t)$ and the second residue $R_2(t)$. The process is continued until either the residue of the intrinsic mode function becomes less than a predetermined small number of the residue becomes nonoscillatory. The resulting IMFs are shown in FIGS. 16(a)–(p).

If the process takes n steps, the original signal $X(t)$ can be represented as follows:

$$X(t)=C_1(t)+C_2(t)+\ldots C_n(t) \qquad 24$$

The last term $C_n(t)$ represents the nonoscillatory mean trend of the signal. As such, important information exists in this mean trend particularly for certain types of data.

The other terms C_{n-1}, C_{n-2}, etc represent the oscillatory portion of the mean trend. Each term represents a different spectral portion of the mean trend.

With the IMFs in hand, further processing is possible. One such processing technique is filtering in which certain IMFs are combined or summed while leaving out other IMFs.

For example, a low-pass filter or low-frequency representation $M_k(t)$ of the mean trend of the signal $X(t)$ can be generated as follows:

$$M_k(t)=C_k+C_{k+1}+\ldots C_n \qquad 25$$

where $2<k<n$

The lower the value of k, the more oscillations $M_k(t)$ contains. By adjusting the values of k and n, a variety of filters may be constructed each having desired characteristic (s).

Representing the original data, e.g. of FIG. 15(a), with an analytic expression is not always possible or practicable. The invention, however, offers a technique of accurately representing the physically meaningful portions of a signal with an analytic function.

For example, when one examines the indicial functions that represent the changes of the signal $X(t)$ in response to step changes in oxygen concentration, one looks for changes in the mean treads $M_n(t)$, $M_{n-1}(t)$, and . . . with respect to changes in O_2 level. To do this, it is helpful to fit the experimental result on $M_k(t)$ with an analytic function. Particularly, if one takes the origin of time t=0 at the time of imposing a step change in oxygen concentration, the resulting change in the mean blood pressure in response thereto for t>0 can be represented as:

$$M_k(t)=A_1+A_2e^{-12^t}+A_3te^{-13^t}+A_mte^{-1}mt \qquad 26$$

where constants A_1, . . . A_m and $_{-2}$, . . . $_{-m}$ may be determined with a conventional least-squares estimation technique.

The above estimation is greatly affected by the choices of k and m. By appropriately selecting k and m, different degrees of detail can be presented.

In reality, one can only approximate a step function change of O_2. The dropping step of FIG. 15(d) and the rising step of FIG. 15(c) cause different changes in blood pressure. Hence, two different empirical functions are utilized. For the first case, the equation is:

$$M_k(t) = \tag{27}$$

$$p_m(t_0)\left[1 + A_1(t - t_0)e^{-1\frac{t-t_0}{T_1}} + A_2(t - t_0)e^{-1\frac{t-t_0}{T_2}} + A_3\left[1 - e^{-l_3\frac{t-t_0}{T_3}}\right]\right],$$

$$\text{for } t_0 \leq t$$

for $t_0 \leq t$, where t_0 is the instant of time when O_2 concentration drops suddenly, and $p_m(t_0)$ is the measured value of $M_k(t_0)$. For the second case the following equation appears good enough:

$$M_k(t) = p_m(t_1)[1 + A_4[e^{-1(t-t_1)} - 1]], \text{ for } t_1 \leq t, \tag{29}$$

where t_1 is the time when the O_2 concentration increases suddenly. The variables A_1, A_2, A_3 and A_4 are dimensionless.

Taking the analysis into the spectral domain, the invention offers further tools. As defined above, the Hilbert transform of $X(t)$ is $Y(t)$. Hilbert has shown that the complex variable $Z(t) = X(t) + iY(t)$ is an analytic function of t and can be written in polar coordinates as $a(t)\exp\{i\_(t)\}$, thus defining the amplitude $a(t)$ and phase angle $\_(t)$.

As further defined above, an instantaneous frequency $\_(t)$ can be calculated from the derivative of $\_(t)$ with respect to time. Knowing the amplitude $a(t)$ and frequency $\_(t)$ as functions of time, one can plot contour maps as shown in FIG. 23(a). The contour maps of the amplitude as functions of frequency and time are called the Hilbert amplitude spectrum, $H(\_,t)$.

The vanishing of the local means of the IMF's $C_1, \ldots C_{n-1}$ is a very important result because the amplitude $a(t)$ and phase angle $\_(t)$ of the Hilbert spectrum are sensitive to the local means.

Using the definition above $\{M_k(t) = C_k + C_{k+1} + \ldots C_n,$ where $2 < k < n\}$ to represent the mean trend, the invention further defines the corresponding sum as follows:

$$X_k(t) = C_1(t) + C_2(t) + \ldots C_k(t) \tag{30}$$

to represent the oscillations about the mean trend $M_{k+1}(t)$. The Hilbert amplitude spectrum of $X_k(t)$ may be designated as $H_k(\_,t)$. The square of $H_k(\_,t)$ represents an oscillatory energy density. Thus, the invention also defines the oscillatory energy about the mean, $E_k(t)$, by an integration over all frequencies as follows:

$$E_k(t) = \int_W H_k^2(w, t)dw \tag{31}$$

The variations of $E_k(t)$ and $M_k(t)$ with oxygen level $O_2(t)$ yield the desired summary of information about the transient response of $X(t)$ to $O_2(t)$.

The oscillatory energy about the mean $E_k(t)$ is quite similar to the instantaneous energy instantaneous energy density, IE, described above. The difference is that the $E_k(t)$ is calculated from a subset (k) of the IMFs while IE is calculated from the full set of IMFs. In other words, the oscillatory energy about the mean is derived from a filtered version of the original signal in which a subset (k) of the IMFs is utilized to construct the filtered signal. This contrasts with instantaneous energy density which is derived from the full set of IMFs.

The oscillatory energy about the mean $E_k(t)$ is another nonstationary stochastic variable that can be treated in the same manner as outlined above. Specifically, the progression from steps 350 to 370 to 120 in FIG. 1(d) provide a process for applying EMD to $E_k(t)$ to generate another set of IMFs. The IMFs thus generated may be further processed according to the techniques disclosed herein. The process may be repeated as many times as desired as indicated by the loop in FIG. 1(d).

From the IMF's shown in FIGS. 16(a)–(p), one can compute the mean trend functions M_k from equation 25 set forth above. The results are shown in FIGS. 17(a)–(f) which are the main results to be used for the indicial response determination.

The next step (360 FIG. 1(d)) is to generate analytic functions or otherwise perform a curve fitting process. This is a chief advantage of the present invention because for many data sets, such as the blood pressure data in FIGS. 15(a)–(d), deriving analytic functions is impossible. By applying the inventive methodologies, however, deriving such analytic functions is possible. Specifically, the mean trend data extracts relevant data and removes unnecessary or irrelevant data thereby producing a simplified data set for which it is possible to derive an analytic function. The degree of simplification and extraction can be controlled by properly selecting which IMF components are included in the mean trend data.

The analytic representation of the indicial response functions of FIGS. 17(a)–(f) by equations 26–28 was done with a least-squares error method. The results are shown in FIGS. 18(a)–(c) and 19(a)–(c). These are the primary results of interest for tissue engineering analysis and design.

Specifically, the results shown in FIGS. 18(a)–(c) and 19(a)–(c) provide indicial response functions of the mean pulmonary blood pressure at the arterial trunk in response to a step decrease or a step increase in breathing gas oxygen concentration. A similar analysis can be directed toward the oscillation modes defined by equation 29.

The oscillation modes for $k=1, 2, 4$, and 6 are respectively shown in FIGS. 20(a)–(d). The signals shown in these figures contain a lot of information that needs to be extracted into simple, understandable terms. This can be done by applying the Hilbert transform and generating the Hilbert spectrum. The results are shown in FIGS. 21, 22(a)–(h), 23(a)–(b). Specifically, FIG. 21 is a plot of the oscillatory energy defined by equation 29 as a function of time. This spectrum is analyzed as a nonstationary random signal by the invention to resolve it into oscillatory IMF modes and mean trend functions $M_k(t)$, with the associated analytic functional representation as illustrated in FIGS. 22(a)–(h) for $k=9, 10, \ldots$, and 16.

To further analyze the data, the results of calculating the instantaneous frequency and amplitude of oscillations of the pressure signal $X(t)$, made possible by the Hilbert transformation, can be plotted against time in a three-dimensional manner as sown in FIG. 23(a). The same can be plotted two-dimensionally on the plane of time and frequency by the contour lines of equal amplitudes, as show in FIG. 23(b).

FIG. 24 is a conventional FFT amplitude spectrum of the pressure signal in 1-min segments under the assumption that the process is stationary in each segment. This conventional spectrum provides a basis of comparison against the Hilbert spectra generated by the invention.

As mentioned above, Fourier analysis is based on the hypothesis of segmental stationary random oscillation, the principle of linear superposition of sine waves, and a global

average of waveform convolution over each time segment. The HHT is based on the hypothesis of nonstationary processes, the principle of linear superposition of nonlinear IMF's, and local determinations of amplitude and frequency (through differentiation rather than convolution) of each IMF. In terms of the IMF, the first k modes can be added together the represent oscillations about the mean trend $M_{k+1}(t)$.

The Fourier series cannot represent time variation of a nonstationary signal and does not have a property or capability to separate a signal into two parts, one part representing a mean trend while the other part represents oscillations about the mean. The number of intrinsic modes, n, is finite: in general, $n < \log_2 N$, where N is the total number of data points. The number of harmonics in Fourier analysis is N/2.

Comparison of the Hilbert and Fourier spectra respectively shown in FIGS. 23(b) and 24 shows that both spectra display a major frequency component at approximate 5 Hz where the energy is concentrated. This is close to the heart rate of the rat subject. This rate decreases when the oxygen concentration is decreased. These two spectra are in different vertical scales. The Hilbert spectrum contains no energy with frequency>10 Hz, and it also has fewer yet more diffused frequency bands that the Fourier spectrum. This is because the Hilbert spectrum gives only the global mean. The mean values certainly will show less variations.

The Fourier spectrum of FIG. 24 contains more frequency bands because any deviation of the waveform from the basic sinusoidal harmonic will result in strong higher harmonics whereas the Hilbert spectrum allows variation of instantaneous frequencies, hence the fuzzy spread in the frequencies. This calls attention to the fact that the heart rate is also a stochastic variable, which could be studied by the inventive methods disclosed herein. The strong higher harmonic band with frequency>10 Hz in the Fourier spectrum is probably spurious. In other words, the Fourier components above 10 Hz (staring at around 15 Hz at time 0) have no physical meaning.

In contrast, the inventive results of FIG. 23(b) show a rich variety of low frequency components thereby conveying important information not available in the conventional result of FIG. 24. Thus, the invention is capable of more a more accurate physical representation of the underlying phenomenon than conventional Fourier-based analysis. Moreover, the clarity of the set of mean trends and the corresponding set of the oscillations is a unique contribution of the invention.

Thus, the invention provides useful tools for concisely and precisely describing the affect of changing one variable on another variable. The effect on blood pressure caused by changing the O_2 concentration provides a good illustrative example of these tools.

To further illustrate the broad applicability of the invention and its potential as a diagnostic tool, an abnormal condition (disease) was chosen. The particular abnormal condition chosen for illustration purposes is sleep apnea. This is a common condition in which the airway is temporarily obstructed during sleep causing the subject to eventually gasp for air. To study this condition, heart rate data was taken from a subject as shown in FIG. 25. More accurately, the data is a measure of the time interval between pulses (pulse interval) of the heart plotted against time. Thus, small values indicate a fast heart rate and large values indicate a slow heart rate. This data includes a normal data section and an abnormal data section, both of which are labelled.

As will be shown in detail below, the invention provides powerful tools for studying this condition. Although the

results to date are preliminary, they do indicate that the invention is capable of diagnosing this condition. Of course, further studies are necessary before a reliable diagnosis can be made using the invention.

To improve the analysis, the data of FIG. 25 was processed to make the spacing between data points even. Furthermore, a section of the normal data (no apnea episode) was extracted for separate analysis. A blowup of the normal data having even spacing between data points is shown in FIG. 26. A further blowup is shown in FIG. 28.

The data of FIG. 26 was then subjected to the inventive EMD to produce eight IMFs, C_1-C_8, as shown in FIGS. 27(a)-(h). The resulting IMFs provide important, and heretofore unavailable, data to a person attempting to gain a better understanding of sleep apnea particularly when these normal IMFs are compared against abnormal (apnea) IMFs.

The IMFs of FIGS. 27(a)-(h) were then utilized to construct the Hilbert Spectrum shown in FIG. 29. This Hilbert Spectrum provides additional information about normal sleep patterns and how they are reflected in the subject's heart rate.

To continue the analysis, an abnormal data section (FIG. 30) was subjected to the inventive EMD to produce eight IMFs, C_1-C_8, as shown in FIGS. 31(a)-(h). The third IMF C_3 is believed to be particularly important in analyzing and perhaps even diagnosing sleep apnea.

FIG. 32 is a blowup of the abnormal data of FIG. 30. The abnormality is apparent around second 5005 where the airway is temporarily blocked thereby causing the heart to race in an effort to extract the diminishing oxygen still available in the lungs. The sharp spike just after second 5005 illustrates the gasping breath and the resulting sharp increase in the pulse rate interval (corresponding to a decrease in the heart rate).

FIG. 33 is the resulting Hilbert Spectrum of the abnormal data of FIG. 30. Comparisons between the normal Hilbert Spectrum of FIG. 39 with the abnormal Hilbert Spectrum of FIG. 30 yields further information for studying and understanding sleep apnea.

To further illustrate the wide applicability of the invention, it was applied to study epilepsy. Epileptic seizures occur when the brain's neurons fire in synchronism. The sequence starts from an epicenter and then propagates to an entire hemisphere of the brain. Even at this stage, the patient can still function in a relatively normal manner. When the synchronization propagates to both hemispheres of the brain, then the patient will suffer a seizure.

To study epilepsy, heart beat rate data was chosen. FIG. 34 shows heart rate data (beats per minute) that was collected from a patient before, during and after a seizure. The time 0 is the starting point of the seizure.

FIGS. 35(a)-(i) show the IMF generated by the inventive EMD method applied to the FIG. 34 data. These IMFs reveal important and physiologically meaningful data not available to a researcher before the application of the present invention.

FIGS. 36(a) and (c) illustrate two IMFs (modes 1 and 3) which may be utilized to correlate the data and study epilepsy. These IMFs are shown in alignment with the original heart rate data (FIG. 36(b)). As is apparent from these figures, the first and third have components that appear to correlate with the onset of a seizure.

The Hilbert Spectrum of FIG. 37 was generated by the invention from the data of FIG. 34. The trace along 18 to 20 cycles/min correlates to time 0, the start of the seizure. To compare the invention against conventional techniques, a corresponding Wavelet Spectrum is shown in FIG. 38.

37

38

As can be seen by the above illustrative examples, the invention is a powerful tool capable of analyzing data and producing new, quantitative measurements that enhance the understanding of the underlying phenomenon producing the data. The IMFs themselves are important results of the inventive data analysis techniques and permit one to gain a level of understanding not possible with conventional techniques.

Although blood pressure, heart pulse interval, and heart rate data provide good examples of the invention's applicability to biological signals, it is to be understood that a large variety of other biological signals may be processed by the invention to gain a better understanding of the underlying biological process(es). Other examples include plethysmogram signals, electro-encephalogram (EEG) signals and temperature signals. Furthermore, the signal need not be taken from a living body and include in vitro studies such as current flow across membranes (patch clamping), fluorescence in confocal microscopy, spectroscopic signals, etc.

As mentioned above, the invention is also not limited to biological signal processing and includes the full range of real-world, data representative of an electrical, chemical, mechanical, optical, geophysical processes or phenomenon or combinations thereof all of which fall under the term "physical signal" as it was defined in the parent application.

Even further, the invention provides tools for studying the influence of one variable on another variable in a multi-variable system. The effect of hypoxia on blood pressure provides one illustrative example of this analysis. However, it is to be understood that any physical signal in a process or phenomenon having multiple variables of which the physical signal is one can be analyzed with the invention.

Alternative Embodiments of Biological Signal Analysis

As described above, the invention constructs upper and lower envelopes **20, 30** with a cubic spline in steps **210** and **230**, respectively and in step **560**. This cubic spline fitting, however, has both overshoot and undershoot problems. These problems can be alleviated by using sore sophisticated spline methods, such as the taut spline in which the tension of the spline curve can be adjusted.

Another alternative is higher-order spline fitting. Although such higher-order spline fitting may be more accurate, it will, however, introduce more inflection points or extrema, and consume more computing time. Therefore, it is not recommended as a standard operation. Only in special cases, it may be tried.

As the spline fitting procedure is time consuming, more efficient methods can be devised by using simple mean values of successive extrema instead of computing the envelope-mean. In this way, only one spline fitting is required rather than two. Although this alternative is easier and faster to implement, the shortcomings are more severe amplitude averaging effects when the neighboring extrema are of different magnitudes. The successive-mean method will have a stronger forcing to reach uniform amplitudes, in which the true physics associated with amplitude will be destroyed. Therefore, the successive-mean method should only be applied where the amplitudes of the physical signal components are constants.

Either the envelope mean or the successive mean method, when applied with the requirement of absolute symmetry, will produce the absurd result of uniform amplitude IMF's. Therefore, the criteria in the Sifting Process should be chosen judiciously. One should avoid too stringent a criterion that we would get uniform amplitude IMF's. On the other hand, one should also avoid too loose a criterion that would produce components deviating too much from IMF's.

It is well known that the most serious problem of spline fitting is at the ends, where cubic splines can have wide swings if left unattended. As an alternative, the invention may utilize a method of adding characteristic waves at the ends of the data span. This confines the large swings successfully.

The method of adding characteristic waves is not conventional. In contrast, the conventional window that is often applied to Fourier transform data results in loss of useful data. To avoid this data loss and to confine swings at the ends of the data span, the invention extends the data beyond the actual data span by adding three additional characteristic waves at each end of the data span.

The characteristic waves are defined by the last wave within the data span at the end of the data span. In other words, a characteristic wave is added to each end of the data span having an amplitude and period matching the last wave within the data span. This characteristic wave is a sinusoidal waveform that is extended three sinusoidal wave periods beyond the data span at each end. This process is repeated at the other end of the data span. In this way, spline fitting at the end of the data span, which can otherwise have a wide swing, is confined. In other words, by adding the extra characteristic waves at the ends beyond the data span, the spline curve will be tied down so that it will not have wild or excessive swings that would otherwise corrupt the data processing and analysis that utilizes these cubic splines.

Other than the spline fitting, the Hilbert transform may also have end effects. Because the first and the last points of the data are usually of different values, the Fourier transform will introduce additional components to bridge over the difference resulting in the well-known Gibbs phenomena. To eliminate it in the Fourier transform, various windows have been adopted (see, for example, Brigham, 1974, "The fast Fourier Transform", Prentice-Hall, Englewood Cliffs, N.J.).

Instead of a window which will eliminate some useful data at the end, the invention again adds two characteristic waves at either end. These waves all start from zero at the beginning, and end at zero at the end. Thus, the annoying Gibbs phenomena are greatly reduced.

Still further, the Hilbert transform needs over-sampled data to define the instantaneous frequency precisely. In Fourier analysis, the Nyquist frequency is defined by two points per wave, but the frequency is defined for a wave covering the whole data span. In the invention, the instantaneous frequency is defined through a differentiation process, and thus more data points will help defining the frequency more accurately. Based on the inventor's experience, a minimum number of data points to define a frequency is five (or four _t's). The lack of fine time steps can be alleviated through interpolating more points between available data. As a spline interpretation would also not create nor annihilate scales, it can also be used for the interpolation when the data is very jagged from under-sampled data. The smoothed data though have a longer length and are sometimes easier to process. The interpolation may give better frequency definition.

Acoustical Signal Analysis

Referring to FIG. **39**, an application of the invention will be to analyze the acoustical signal through the Hilbert-Huang Transformation (HHT), which consists of the Empirical Decomposition Method (EMD) and the Hilbert Spectral Analysis (HSA). Essentially, the signal will be decomposed into the Intrinsic Mode Function Components (IMFs). This decomposition is outlined in Huang's publications and patents. Once the invention decomposes the acoustic signal into

its constituting components, all the operations can be performed on these components.

Speech Analysis and Speaker Identification:

Referring to FIGS. 41 and 42, we will first discuss the way HHT analyzes speech. Then, based on the analysis results, we can discuss how to identify a speaker, i.e., to single out one speaker from others. Let us start with the following example.

Two speakers both said 'Halloo,' and the acoustic signals are digitalized at 44100 Hz with 16 bits resolution as shown in FIG. 45a for a male speaker, and b for a female speaker. The EMD decomposed IMF components are given in FIG. 46a and b. Each signal consists of 12 components, and each component consists of 30,000 points. They are plotted with the same vertical scale in FIG. 46. As the IMFs are the results of scale separation of the original signals, each component represents different time scales, which can be translated to the frequency ranges. It should be pointed out that this frequency separation is not in the Fourier sense, for the signals separated in scale could be nonlinear; therefore, it can contain both sub and super harmonic components in Fourier frequency space. Consequently, the decomposition is not necessarily narrow banded in Fourier sense, but, as explained in Huang (1999), this decomposition preserves all the nonlinear properties of the signal. One can immediately see from these figures that the two decompositions contain very different IMFs. For the signal given in FIG. 45a, the decomposition is given in FIG. 46a. There, we have very regularly periodic signals for all the IMFs. Comparing this with the corresponding results in FIG. 46b from the second speaker, we can see the differences. The second speaker does not have the regularity in the time scale. In the second speaker's results, we can see that the strongest signal is the third component, but for the first speaker the corresponding component is relatively weak. By the time we come to the 6th component of the first speaker, the component is still very strong, but there is almost no more signal in the second speaker's result in FIG. 46b. The richness of the low frequency components for the male speaker, and high frequency for the female speaker should be expected. This difference in frequency components can be seen vividly in the Hilbert spectral analysis as follows.

From these components, we can construct the Hilbert Spectra as given in FIGS. 47a and b, in which time-frequency-energy representation of the signals are presented, with the corresponding narrow banded Fourier based spectrograms given in FIGS. 48a and b constructed with 1028 sampling points in the window, which corresponding to the window width of 43 Hz. In these figure, the total frequency range is represented by 200 bins to cover the frequency range from 0 to 10,000 Hz. In FIG. 47a, we can identify a clear consistent low frequency energy concentration near 100 to 200 Hz, but there is no equivalent energy at this low a frequency in FIG. 47b. To examine the results in more details, we have given the Hilbert spectra of the cases presented in FIG. 47a and b in FIGS. 49a and b, in which the same 200 frequency bins are reassigned to cover the frequency range from 0 to 1,000 Hz. In these new and more detailed presentations of the signals, we can see that the male voice represented by the Hilbert spectrum by FIG. 49a contains a strong modulated energy below or around 100 Hz, while the female voice represented by the Hilbert Spectrum given in FIG. 49b contains the most energetic modulated components above 100 Hz. Furthermore, the female voice also contains wider range of high frequency energy distributions. In FIG. 49, we can see the regular spaced vertical striations in the Hilbert spectra, which are the indications of

the resonant beat period from the speaker's vocal cavities. This beat pattern can be easily seen from the IMF components, which can serve as crucial data for speaker identification too.

In order to show the unique properties of the HHT analysis results, we decided to examine the male voice in more details as follows: First, we give the spectrum it wide banded Fourier spectrogram as in FIG. 50, constructed from 128 sampled points in a window, which corresponds to a wide banded spectrogram with band width of 344.5 Hz. Clearly neither the narrow nor the wide banded spectrograms give us any detailed information. This is shown even more clearly if we present the detailed views in FIGS. 51a and b, in which the spectrograms are plotted for the frequency range between 0 to 1000 Hz, corresponding to the condition represented in FIG. 49a. The comparison is clear, while the Fourier based spectrograms could give us some qualitative information, the results is so crude that they cannot be used as a based for any serious acoustic signal analysis.

In order to emphasize the superior performance of the HHT, we also process the same data with Morlet Wavelet analysis, a continuous Wavelet analysis. The result is given in FIGS. 52a and b. These figures clearly identify the different time-energy-frequency distribution characteristics between male and female speakers, the strong high frequency energy concentration for female and the strong low frequency energy concentration for the male speaker. This clear difference aside, the Wavelet results really lacks the detailed analysis ability of HHT. When on comparing the pair of FIGS. 52a and b with those in FIGS. 49a and b, one can immediately see the lack of time-frequency resolution of the Wavelet results. These differences have been discussed in great details by Huang et al (1998, 1999).

To illustrate the lack of resolution in the Wavelet analysis further, we can smooth the Hilbert Spectrum result given in FIG. 49a by filters in 100 and 200 time-steps. The results are given in FIGS. 9a and b respectively. By the time we smoothed the results with filter width of 200 points, the results are clearly much smoother, but we will never obtained the kind of overly smoothed result given by the Wavelet, for the two analysis methods are very different: The Hilbert spectrum gives the genuine instantaneous frequency based on an adaptive basis totally independent of the kind of a priori basis adopted in the Wavelet analysis. Furthermore, the components in HHT are almost orthogonal, which gives no energy leakage among different modes. The continuous Wavelet, on the other hand, has a non-orthogonal representation. The energy leakage among different frequency bands, and the over redundancy have rendered the method to only extract qualitative information.

The detailed time-frequency-energy presentation gives us many more reference points for speaker identification and verification. The continuous modulation of frequency and energy will provide more reference points to separate one speaker from the other. HHT analysis could also identify peculiarities of one speaker from another. All these details just cannot be found in any other analysis method.

To summarize the material discussed in this section, we will first give a flow chat of speech analysis: Speech containing various phonemes will be recorded, and the signal will pass through the sifting process to get the IMFs. In this step, the critical task is to identify the phonemes and formants (Hilbert Spectra) for each sound element with modulations in both frequency and amplitude as functions of time. Databases will be build from the recorded speech data. Such analysis is not available from any other method of analysis.

There are two types of speaker identifications: to verify that the speaker is a known person, and to separate the speech from different speakers. The essential steps for both tasks are as follows:

To verify a known speaker, we have to have the database of the targeted speaker built ahead of time. The database can be built following the steps listed in the speech analysis. To separate speakers from one another, we can record the speakers and analyze their speech as listed in the speech analysis part. Once the analyses are done, we can compare the details of the Hilbert spectra in their over all and the details of the spectral patterns, including the beat periods for the same phoneme and formant. Given the super temporal and frequency resolutions of the Hilbert spectra, the Hilbert spectra actually give a much better formant representation. With so many details in the Hilbert spectra, we can use many features as references. In the example given in this section, the overall pattern is sufficient.

For Speech Recognition:

Speech analysis such as speech recognition as shown in FIG. **42**, the goal is not to identify a specific speaker. Rather the task is to identify the meaning of an utterance so that the commands contained in the speech can be translated into actions as in man-machine interaction. Because, the speakers can have a variety of accents, the approach is to obtain a limited number of uniquely identifying parameters to define the meaning of the utterance, rather than to obtain more and more detailed descriptions of the sound. In such applications, two approaches can be adopted for the HHT.

The first is by smoothing as in the cases of FIGS. **53***a* and *b*. More smoothing can be obtained by time-frequency averaging, rather than just time smoothing described here. The Laplacian filter is readily available. However, for most acoustic signal applications, there are lot more time steps than possible frequency bins. Therefore, time-wise smoothing is a more logical approach. Other than smoothing to obtain the time-frequency-energy pattern as means for recognition, the formants identification is also a powerful method. Here we will discuss this approach as follows:

Formants by definition are the few characteristic time-frequency distribution of the maximum energy concentration in a speech section analyzed with filter bank or spectrogram methods (see, for example, Rabiner and Schafer, 1978). Conventionally, to make the number of components manageable, the formants are defined from the low frequency components only. The ideal of formant is to sharpen the poor time-frequency analysis in the spectrograms. Similar approaches have also been explored by Carmona et al (1997) using Wavelet analysis. As discussed by Huang et al (1998, 1999), the inherited poor time-frequency resolution cannot be overcome by post analysis smoothing or selection. The only alternative is to change to a genuine instantaneous frequency analysis as in the HHT, which in form and in substance is a much better formant representation of the speech. Let us examine an example. If we take the beginning of the sound for 'Halloo' signal, and analyze it with HHT, the result is given in FIGS. **54***a* and *b*. This is actually the consonant part of the 'h' sound. These figures, in fact, are simply obtained by the enlargement of the same section of the figure of **47***a* and **49***a*. One of the great advantages of HHT is that the frequency is defined by differentiation; therefore, we can breakthrough the limitation of the uncertainty principle and expanding the results to the limit of the sampling rate.

FIG. **54***a* is the expansion of FIG. **47***a*, while FIG. **54***b* is the expansion of FIG. **49***a*, both with the original signal plotted. Here we have the lower frequency emphasized. The

continuous lines of the time-frequency curve are in form and substance a better formant. But they are indeed better defined than the conventional formants, for they also included the nonlinear deformation of the signal rather than assigning all the waveform deformation to harmonics. As the frequency is defined by differentiation, the information provided here is might be too much. If that is the case, the frequency variation can be simplified by smoothing either through filtering, or defined as differentiation with larger time steps.

The essential steps for this task can be summarized in the following flow chart: Again, we have to establish the database consisting of templates of the key phonemes and formants. In this case, the templates will be constructed from the averaged value of a large number of different speakers. In addition to the averaged value, the range of scattering from the differences will also be recorded. When a new speech is given, the record will be analyzed as in the previous part to have the speech breakdown to phonemes and formants. Each sound element will be compared with the existing templates and identified. The identified sound will be put together, and translated into sentences. Grammar check and corrections should be made here to guarantee that the speech makes sense. The final text will then be translated into instruction for further actions, such as an order to a servomechanism.

Speech Synthesis:

Referring to FIG. **40** the speech synthesis problem has been review recently by Breen (1992). Basically, the synthesis is based on a phoneme base. Having established the analysis part, we can take sections of the sound record as phonemes and build a database for it. This phoneme database can then be used to synthesize speech. The advantage of this phoneme base has the unique advantage over the conventional one in that no local stationary assumption has been evoked. Therefore, the frequency is continuously modulating, which can make the final result sounds more naturally.

The detailed time-frequency-energy distribution results given by the HHT also give us a means to synthesize sound with higher fidelity, for the frequency modulation can be truly simulated by the instantaneous values in frequency and energy. This overcomes one of the fundamental shortcomings of the Fourier based analysis, in which a local stationary assumption is to be invoked. This piecewise stationary analysis simply cannot closely simulate the constantly modulated frequency modulation as shown in the Hilbert spectral analysis results.

Also from FIGS. **54***a* and *b*, we can define the beginning of the speech clearly and un-ambiguously. The separation of syllables has also been a hard problem in speech analysis and synthesis.

A key to the potential value of our approach is that HHT gives time varying modulation of frequency and amplitude. The essential steps are as follows:

Based on the database of all the phonemes and formants, we can translate the given text to the time varying IMF components from the database. This translated data is checked again with prosody patterns, and synthesized into sound.

Speech Quality Enhancement:

Referring to FIG. **43** one of the most important applications of acoustic signal analysis is to enhance the speech quality. For example, during the utterance of the word, 'Halloo,' a noise simulated by the drop of a spoon on the table was also heard. How to separate this noise from the speech signal is an important application. Traditionally, the cleaning up of a speech is carried through filtering But

Fourier based filtering has inherited shortcomings: Localized signal in time is broad in frequency; localized signal in frequency is broad in time. Furthermore, as discussed above, the generating mechanisms of most consonant sound are nonlinear; therefore, most of the consonants, the information carrying sounds, have a lot of sub as well as super harmonics. We cannot remove certain frequency without altering the harmonics of other sound signal, which will down-grade the filtered result through loss of the harmonics. Here we propose the filtering through EMD. The filtering strictly by scale, the filtering here has to be implemented jointly in time-frequency space.

Let us examining the signal of the sound, 'Halloo+ding,' given as the top line in FIG. 55. This signal is decomposed through EMD and we obtained the IMF components as given in FIGS. 56a and b. FIG. 56a gives all the components with the uniform vertical scale, while FIG. 56b gives the same information with the vertical scale in each IMF components normalized within its own frame. From FIG. 56a, we can see that the 'ding' sound, occurring slight after 1.5 second mark, has the higher amplitude than any other IMF components. FIG. 56b indicates that at the same time scale components with the 'ding' sound, there are also signals from the 'Halloo,' as indicated by the low level signal in c3. As a result, we cannot simply filter out the high frequency components by eliminating the first three IMFs.

Now we will accesses the performance of the EMD vs Fourier filtering. First, let us implement only the filtering by simply eliminate the first three IMF components. This approach is similar to the thinking of Fourier filtering as discussed by Huang (2000), but the result is not exactly the same. We will design this filtering as the "Fourier" type and discuss the difference later. The result certainly does not show the 'ding' sound any more. But this brute force filtering has also eliminated all the high frequency sound including some higher frequency elements from the 'Halloo' sound. The quality of the sound is degraded considerably. A better way is to eliminate the 'ding' sound in both time and frequency space by selectively removing energy associated with the 'ding' sound. This was achieved by first studying each of the IMF component, identifying the 'ding' sections in each component by their sounds, frequencies, and occurrence time location. Then, selectively removing the sections that contain the 'ding' associated signals from each IMF component. After this is done, and separated 'Halloo' and 'ding' signals are represented by their respective IMF decompositions in FIGS. 57a and b. In FIG. 57a, we can see that there is no longer signal associated with the 'ding' sound, while, in FIG. 57b, we see the clear 'ding' sound stands by itself. The Hilbert spectra for the original and the EMD filtered signals are shown in FIGS. 58a and b. When all the IMF's are summed up, the re-constituted sound signals for 'Halloo' and 'ding' are given in FIG. 55. The sound quality is so much improved over the brute force approach.

Next, we will implement the Fourier based filtering with two 48 degree FIR filters with cutoff frequency set at 2000 and 1000 Hz respectively. The filtered data with a cutoff frequency at 2000 Hz together with the HHT filtered data is given in FIG. 59a. Based on the results plotted here, one can hardly see the difference. But if one pays attention to the relative darkness of the two side-by-side signals, one would see that Hilbert filtered signal is darker, an indication of more high frequency components being retained. Furthermore, if one plays the filtered data back through a speaker, we can immediately tell the difference, for there are still residual 'ding' sound in the Fourier filtered case. We

then filtered the data with a similar filter but with a cutoff frequency set at 1000 Hz. This certainly eliminated the 'ding' sound, but the quality of the sound was very much downgraded.

The Fourier filtering is strictly in frequency space, while EMD filtering is in time scale. By operating in a rigid frequency range, the Fourier filter will not eliminate the bounded super and sub harmonics of the signal outside of the pre-selected filtering frequency range. It, however, will remove the bounded super and sub harmonics of other signal components that happened to fall within this pre-selected frequency range, even if the fundamentals are outside this frequency range. As a result, through Fourier filter is localized in frequency, it actually affects signals of all frequency. In EMD, each IMF could be nonlinear; therefore, each of them is not banded limited in Fourier sense. By eliminating the first three IMF components, we still could have energy at Nyquist frequency in Fourier sense through the existence of the harmonics. So, through filtering by EMD, we can really limit the frequency to the components in time scale rather then in Fourier sense. This time space filtering will not influence the lower frequency components because it would leave the bounded super and sub harmonics of other components untouched, and it will also eliminate all the bounded super and sub harmonics associated with the selected time scale component. We have plotted the results of the Fourier spectra of the EMDT filtered, the "Fourier" type of filtering by eliminating the first three IMF components, the Fourier based FIR with cutoff frequency set at 2000 and 1000 Hz all in FIG. 59b. As we discussed above, only by setting the cutoff frequency at 1000 Hz can we eliminate the 'ding' sound. The spectra showed the difference: Fourier filtering has eliminated all the energy at the high frequency range, but EMD filtering can leave some energy there. The result is a greatly improved voice quality.

Another application for the present invention is on music recording. Using EMD rather then Fourier based filter we can surgically remove the unwanted noises or sound to enhance the final quality of the product. As discussed above, the EMD filter is much more sophisticated than the Fourier based one, for the EMD filter operates with the signal separated by temporal scales; therefore, it can keep track of the events precisely on the temporal axis. This operational principle is totally different from the Fourier based filter, which operates in Frequency space only. As discussed above, the localized signal in frequency space means the signal is uniformly distributed in time space, and localized signal in time space means uniformly distribution in frequency space, it is just impossible to perform the precision surgical-like filter localized both in time and frequency spaces. Furthermore, many of the sound full of harmonics will be degraded no matter which frequency the Fourier filter will remove. In the EMD filter, however, the IMF components retained all the harmonics within a single scale component. The harmonics are represented by the frequency modulation as discussed by Huang et al (1998, 1999) and Huang (1999). Consequently, any filtering operation will not downgrade signal at other scales, which can still retain all its harmonics, if viewed in theFourier sense.

A step in this task is to identify the unwanted sound in its time coordinate. It is impossible to blindly apply this method as in the traditional Fourier analysis to remove high frequency components. If that is the case, then we can also use EMD to remove the highest frequency IMF component. This approach has one unique advantage: By removing the highest frequency component locally or globally, it will not alter the quality of the low frequency sounds, for HHT does not

depend on harmonics to represent nonlinear distortion of the waveforms of sounds. With the additional time tag, we can also implement HHT filter locally and globally with the precision unmatched by Fourier window-based filters.

If the unwanted sound is identified on the time axis, we can zero in the IMFs and test each component, or a combination of several components, to determine which one or which components should be removed. Once the unwanted component (or components) is removed, the cleansed sound could be re-constituted locally.

Acoustical Signal Storage and Transmission:

Acoustical signal storage and transmission might require limited bandwidth or data quantity. Using HHT, one could also reduce the amount of data to represent a given signal with acceptable, albeit reduced, quality. This can be achieved by re-sampling of the IMF components of longer time scales at a lower sampling rate. At the same time, we can also eliminate the short time scale components without too much of a reduction of the signal quality. The result of this re-sampling will enable us to store and transmit a given signal with reduced amount of data.

Data storage is a concern as the HHT is designed to examine the details of the signals. For storage, however, there are steps HHT can help. These steps are described as follows:

From the designed signals, decompose the data into IMFs. Test the recombined IMF without the highest few components to determined the degradation of the sound quality. Within the tolerable range, eliminate the maximum number of high frequency components. Re-sample the data by decimating (i. e., with large sampling steps). This will result in a smaller number of data, but also with slightly down graded sound quality.

Machine Operating Condition Monitoring

Referring to FIG. 44 the acoustic signal emitted from an operating machine can be compared to a kind of speech made by the machine trying to inform us of it condition. A chipped gear, a broken bearing, a cracked shaft, a worn grinder, cutter, or saw blade will all send its distinct sound and particular vibration signals. Such signals can be analyzed much the same way as speech signals discussed above. We will use a grinder as an example.

FIGS. 60a, b and c show the acoustical data from a grinder operating on a hard surface continuously for 95, 120, and 200 hours designed as TR095, TR120 and TR200 respectively. From the signals, it is hard to say whether the grinder operated at its optimal condition. Of course, the magnitude of the signal representing the loudness of the sound is increasing, but that alone is not a criterion for determining the condition of the grinder, for there might not be a priori record of the grinder at other time periods. Also from the record, we can determine the rotating speed of the grinder. The rotating speed is also changing. For the cases of TR120 and TR200, the rotating speed is 8,000 rpm; while the rotating speed for the case of TR095 is slightly higher at around 14,000 rpm. None of the above characteristics can be used as an indication of the health condition of the grinder.

To determine the health condition of the machine, one has to establish a criterion based on the condition of the machine at the time of monitoring. This can be implemented by HHT analysis. Using HHT analysis, we can first sift the data and obtain their IMF components as shown in FIGS. 61a, b and c. Then, we can also construct the Hilbert spectra from the IMFs as shown in FIGS. 62a, b and c. Here some distinct characteristics appear. When the grinder is sharp and operating smoothly, there is a clear frequency of the acoustic signature as in FIG. 62a. The energy concentrates around the

rotating speed as indicated. As the grinder becomes dull, it happens first to only part of the grinder surface. The noise generated by the dull part causes signal with greater amplitude or energy density. This intermittent acoustic signature occurs proportional to the dull surface of the grinder. This diffusion of acoustical signal can cover a much broader range, but only intermittently. The intermittence, however, is only detectable by the HHT analysis as shown in FIG. 62b. To the ear, the sound might still be continuous. As the wearing progresses, all parts of the grinder becomes dull. At this stage, the acoustical signal will be a diffused signal everywhere as shown in FIG. 62c.

The grinder is but only one example. Other telltale signs from the chipped gear, the broken bearing, the cracked shaft, the worn cutter, or saw blade can be identified in a similar way.

To summarize the steps in machine health monitoring, we can state that this task is similar to speech identification. We can establish the templates database of normal and various abnormal operation conditions from the sound and/or the vibration signal of the machine. To diagnosis a machine, just record the operational sound and/or the vibration signal of the machine at any time to make comparison with the templates in the database to identify the problems. Knowing problems with existing templates, the problem can be identified. For safety sake, whenever the machine sound and/or its vibration signals are not within the tolerable range of the normal condition template, warning signal should be given.

The dependence on the existence of scale for mode definition has a limitation in that the decomposition method cannot separate signals when their frequencies are too close. In this case, there would not be any characteristic scale: therefore, physically they are identical.

Particular Advantages of The Invention

The strength of the EMD method should be reiterated. EMD is built on the idea of identifying the various scales in the data which are quantities of great physical significance. Therefore, in the local extrema and curvature extrema Sifting Processes, orthogonality is not a consideration, but scales are. Since orthogonal decomposition is a characteristic for linear systems, violating this restriction is not a shortcoming but a breakthrough. Therefore, the decomposed IMF's may or may not be orthogonal. As such, this method can be applied to nonlinear data. Though the IMF's in most cases are practically orthogonal, it is a coincidence rather than a requirement of the EMD.

Another advantage of the method is the effective use of all the data representing the physical phenomenon. In the Sifting Processes, the longest scale is defined by the full length of the data. As a result, EMD can define many long period oscillations. As is well known, the Hilbert transform without sifting tends to identify the highest frequency (Boashash, 1992, "Estimating and Interpreting the Instantaneous Frequency of a Signal, Part I: Fundamentals", Proc. IEEE, 80, 520–538.), the extraction of the long period components is indeed a new feature of the EMD.

Finally, though the EMD method will give IMF components, the individual component does not guarantee well-defined physical meaning. This is true for all decompositions, especially for the methods with a priori basis. In most cases, however, the IMF's do carry physical significance. Great caution should be exercised in making such attempts. The rule for interpreting the physical significance of the IMF's is that the scales should be clearly separated. Together with the Hilbert spectrum, the totality of the presentation should give a much more detailed representation of the physical processes than conventional methods.

US 6,862,558 B2

47

The invention being thus described, it will be obvious that the same may be varied in many ways. Such variations are not to be regarded as a departure from the spirit and scope of the invention, and all such modifications as would be obvious to one skilled in the art are intended to be included within the scope of the following claims.

What is claimed is:

1. A computer implemented method of analyzing an acoustical signal, comprising:

inputting the acoustical signal;

extracting a set of intrinsic mode functions from the acoustical signal; and

storing said set of intrinsic mode functions of the acoustical signal; and

identifying a specific acoustical signal;

wherein said specific acoustic signal is identified in said set of intrinsic mode functions.

2. The computer implemented according to claim 1, wherein said specific acoustical signal is noise.

3. The computer implemented method according to claim 1, further comprising:

removing said specific acoustical signal from said set of intrinsic mode functions; and

reconstructing the acoustical signal.

4. A computer implemented method of analyzing an acoustical signal, comprising:

inputting the acoustical signal;

extracting a set of intrinsic mode functions from the acoustical signal;

storing said set of intrinsic mode functions of the acoustical signal;

transforming said set of intrinsic mode functions with a Hilbert transform to generate a Hilbert spectrum;

identifying a specific acoustical signal in the Hilbert spectrum; and

storing the Hilbert spectrum.

5. The computer implemented method according to claim 4, wherein said specific acoustical signal is noise.

6. The computer implemented method according to claim 4, further comprising:

removing said specific acoustical signal from said set of intrinsic mode functions; and

reconstructing the acoustical signal.

7. A computer implemented method of analyzing an acoustical signal, comprising:

inputting a first acoustical signal;

extracting a first set of intrinsic mode functions from the first acoustical signal;

transforming said first set of intrinsic mode functions with a Hilbert transform to generate a first Hilbert spectrum; and

storing said first Hilbert spectrum;

inputting a second acoustical signal;

extracting a second set of intrinsic mode functions from the second acoustical signal;

transforming said second set of intrinsic mode functions with a Hilbert transform to generate a second Hilbert spectrum;

storing said second Hilbert spectrum of the second acoustical signal; and

comparing said first and second Hilbert spectra.

48

8. The computer implemented method according to claim 7, wherein the first acoustical signal is generated from a first human voice source.

9. The computer implemented method according to claim 7, wherein the first acoustical signal is generated from a first human voice source and the second acoustic signal is generated from a second human voice source.

10. The computer implemented method according to claim 7, wherein the step of comparing said first and second Hilbert spectra includes obtaining a correlation coefficient between said Hilbert spectra.

11. The computer implemented method according to claim 7, further comprising:

providing a second Hilbert spectrum; and

comparing said first and second Hilbert spectra.

12. A computer implemented method of analyzing an acoustical signal, comprising:

inputting the acoustical signal;

extracting a set of intrinsic mode functions from the acoustical signal;

storing said set of intrinsic mode functions of the acoustical signal;

identifying a specific acoustical signal;

removing said specific acoustical signal from said set of intrinsic node functions; and

reconstructing the acoustical signal.

13. The computer implemented method according to claim 12, wherein reconstructing the acoustical signal includes summing up said set of intrinsic mode function.

14. A computer implemented method of analyzing an acoustical signal, comprising:

inputting the acoustical signal;

extracting a set of intrinsic mode functions from the acoustical signal;

storing said set of intrinsic mode functions of the acoustical signal;

transforming said set of intrinsic node functions with a Hilbert transform to generate a Hilbert spectrum;

identifying a specific acoustical signal in the Hilbert spectrum;

removing said specific acoustical signal from said set of intrinsic mode functions; and

reconstructing the acoustical signed.

15. The computer implemented method according to claim 14, wherein reconstructing the acoustical signal includes summing up said set of intrinsic mode function.

16. A computer implemented method of analyzing an acoustical signal, comprising:

inputting a first acoustical signal;

extracting a first set of intrinsic mode functions from the first acoustical signal;

transforming said first set of intrinsic mode functions with a Hilbert transform to generate a first Hilbert spectrum: and

storing said first Hilbert spectum;

wherein the first acoustical signal is generated from a first human voice source.

US 6,862,558 B2

49

17. A computer implemented method of analyzing an acoustical signal, comprising:

    inputting a first acoustical signal;

    extracting a first set of intrinsic mode functions from the first acoustical signal;

    transforming said first set of intrinsic mode functions with a Hilbert transform to generate a first Hilbert spectrum;

    storing said first Hilbert spectrum;

    providing a second Hilbert spectrum; and

    comparing said first and second Hilbert spectra.

50

18. The computer implemented method according to claim **17**,

    wherein the step of providing the Hilbert spectrum of the specific acoustical signal includes retrieving said second Hilbert spectrum from a database.

19. The computer implemented method to claim **17**,

    wherein the step of comparing said first and second Hilbert spectra includes obtaining a correlation coefficient between said Hilbert spectra.

* * * * *