2014374349 09 Oct 2017

DAVIES COLLISON CAVE
INTELLECTUAL PROPERTY

Davies Collison Cave Pty Ltd

Patent and Trade Mark Attorneys
Australia and New Zealand

1 Nicholson Street
Melbourne Victoria 3000
Australia

GPO Box 4387
Melbourne Victoria 3001
Australia

T +61 3 9254 2777
F +61 3 9254 2770
E mail@davies.com.au
W www.davies.com.au
ABN 13 613 954 368

9 October 2017

The Commissioner of Patents
WODEN ACT 2606

Examination - Response to an Examiner's Report

DCC Ref: 35249178/WJP/SJD

Re: Massachusetts Institute of Technology
Australian Patent Application No. 2014374349
"Using correlation structure of speech dynamics to detect neurological changes"

Dear Commissioner,

This submission is being filed in response to the First Examination Report dated 4 November 2016.

We ***attach*** a First Statement of Proposed Amendments.

We explicitly withdraw any request for postponement of acceptance previously made.

Amendments

We request leave to amend the complete specification in accordance with the accompanying Statement of Proposed Amendments. The amendments do not result in the specification claiming or disclosing matter beyond that disclosed in the application as filed.

Claims 4, 9, 11, 21 and 24-35 are proposed to be deleted, without prejudice.

Claim 1 is proposed to me amended to replace "at least one speech related variable" with "at least one *vocal tract representation*". Similar amendments are also proposed for claims 2-6, 11, 12 and 15-17. Support for the proposed amendments may be found at least at page 12, lines 4-8.

Claims 1 and 15 are also proposed to be amended to introduce the feature of processing an input representing a subject's speech, resulting in at least one vocal tract representation. Support for these amendments may be found at least in Example 1, in particular, sections 2 and 3.

Claims 4, 5, 18 and 19 are also proposed to be amended to correct a typographical error in the word "Coefficient". Similar amendments are also proposed throughout the description.

AUSTRALIA
NEW ZEALAND
SINGAPORE
ASIA PACIFIC

The nature and location of the proposed amendments are believed to be clear from the marked up pages provided.

Amendments are also being proposed to bring the specification into better conformity with Australian practice.

Remarks

With reference to item 1 of the Examination Report:

The Examiner has objected to claims 1-35 as lacking novelty and inventive step in view of the disclosures of prior art documents D1. We respectfully disagree and submit that the claims as amended are both novel and inventive over the prior art.

D1 describes a feature analysis approach that utilises multiple electroencephalogram (EEG) signals to predict seizures. Seizures are caused by irregular electrical activity in the brain. Therefore, D1 provides a predictive algorithm that detects changes in brain dynamics as measured according to the spatiotemporal correlation structure of EEG signals.

By contrast, the presently claimed invention is directed to a method for assessing a condition in a subject based upon at least one vocal tract representation. Vocal tract representations are aspects of speech production, which are used according to the present invention to assess changes in coordination in vocal tract spectral shape dynamics (*see* specification, page 12, lines 2-8). There is no teaching or suggestion in D1 that vocal tract representations may be used to predict seizures. In fact, the focus of D1 is to assess the performance of the algorithm in patients with medically intractable focal epilepsy. Therefore, we submit that the person skilled in the art would not have been directly led to the presently claimed invention based on the disclosure of D1.

For at least these reasons, we respectfully submit that the present claims are both novel and inventive over the prior art.

It is submitted that the application is now in order for acceptance.

Yours respectfully,
DAVIES COLLISON CAVE PTY LTD

Bill Pickering
Principal
bpickering@davies.com.au

AUSTRALIAN PATENT APPLICATION NO. 2014374349

Massachusetts Institute of Technology

FIRST STATEMENT OF PROPOSED AMENDMENTS

1. Cancel pages 2, 4, 8 and 9 currently on file and substitute with new pages 2, 2a, 4, 8 and 9 (as enclosed with a copy of the original pages showing the nature/location of the proposed amendments).

2. Cancel claim pages 26-30 currently on file and substitute with new claim pages 26-28 (as enclosed with a copy of the original claim pages showing the nature/location of the proposed amendments).

9 October 2017

speech-related variable in a subject; extracting a channel-delay correlation structure of the at least one speech-related variable; and generating an assessment of a condition of the subject, based on the correlation structure of the at least one speech-related variable.

[0006] In another embodiment, the present invention is a system for assessing a condition in a subject. The system comprises a speech-related variable measuring unit that measures at least one speech-related variable in a subject; a channel-delay correlation structure extractor that extracts a correlation structure of the at least one speech-related variable; and an assessment generator that generates an assessment of a condition in the subject based on the correlation structure of the at least one speech-related variable.

[006A] In another embodiment, the present invention provides a method of assessing a condition in a subject, the method comprising: processing an input representing a subject's speech resulting in at least one vocal tract representation of the subject; extracting, by a structure extractor, a channel-delay correlation structure of the at least one vocal tract representation; and generating an assessment of a condition of the subject, based on the correlation structure of the at least one vocal tract representation.

[006B] In another embodiment, the present invention provides a system for assessing a condition in a subject, the system comprising: a processing unit that processes an input representing a subject's speech resulting in at least one vocal tract representation of the subject; a channel-delay correlation structure extractor that extracts a correlation structure of the at least one vocal tract representation; and an assessment generator that generates an assessment of a condition in the subject based on the correlation structure of the at least one vocal tract representation.

[006C] Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

[006D] The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

[0007] The methods and the systems described herein are advantageously language-independent. Additional advantages include channel-independence as the methods and systems disclosed herein employ data features that do not change with noise or power.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[0009] The foregoing will be apparent from the following more particular description of example embodiments of the invention, as illustrated in the accompanying drawings in which like reference characters refer to the same parts throughout the different views. The drawings are not necessarily to scale, emphasis instead being placed upon illustrating embodiments of the present invention.

[0010] Figure 1 is a color-coded two-dimensional plot showing an example of a channel-delay correlation matrix computed from formant tracks from a healthy subject (top panel) and from a severely depressed subject (bottom panel).

[0011] Figure 2 is a plot of eigenvalues as a function of eignevalue rank, with eigenvalues ordered from largest to smallest (*i.e.* an eigenspectra), derived from formant channel-delay matrices shown in Figure 1.

DETAILED DESCRIPTION OF THE INVENTION

[0019] A description of example embodiments of the invention follows.

[0020] As used herein, the term “speech-related variable” means an anatomical or a physiological characteristic of a subject that can be measured during the subject’s speech and can serve as a basis for generating an assessment of a condition of the subject, as described herein. Examples of speech-related variables include formant frequencies, as defined below, Mel Frequency Cepstral Coefficients (MFCC) and Delta Mel Frequency Cepstral Coefficients (Delta MFCC), as defined below, prosodic characteristics of speech (that is any characteristic of speech that provides information about the timing, intonation, and/or energy), facial features of the speaker, and skin conductance of the speaker.

[0021] Additional examples of speech-related variables include pitch, aspiration, rhythm, tremor, jitter, shimmer, other amplitude- and frequency-modulation functions, as well as their frequency decompositions.

[0022] In some embodiments, certain speech-related variables are referred to herein as “low-level features.” Such low-level features include the following. **Harmonics-to-noise ratio (HNR):** HNR is an estimate of the harmonic component divided by the aspiration component in voiced speech, and can act as a measure of “breathiness” in a voice. It is computed over successive frames (e.g., every 10 ms). Aspiration occurs when turbulence is generated at the vibrating vocal folds.

[0023] **Cepstral peak prominence (CPP):** CPP is defined as the difference, in dB, between the magnitude of the highest peak and the noise floor in the power cepstrum for a time interval of greater than about 2 ms and is computed over successive frames (e.g., every 10 ms). (The cepstrum is defined as the Fourier transform of the log-spectrum.) Several studies have reported strong correlations between CPP and overall dysphonia perception, breathiness, and vocal fold kinematics. **Facial action units (FAUs). FAU represent** measurable differences between facial expressions, and relate to facial features derived from optical video of the face that correspond to muscle movements of the face. The facial action coding system (FACS) quantifies localized changes in facial expression representing facial action units (FAUs) that correspond to distinct muscle movements of the face.

[0042] As used herein, the term “Mel Frequency Cepstral Coefficients” (MFCC) refers to the coefficients that collectively make up a “mel-frequency cepstrum” (MFC), which is a representation of the short-term power spectrum of a sound signal. The term “cepstrum” refers to the result of taking the Inverse Fourier transform (IFT) of the logarithm of the spectrum of a signal. The term “mel” refers to the use of the “mel scale” or similar filterbank by the methods that obtain MFCC. The “mel scale” is a perceptual scale of pitches judged by listeners to be equal in distance from one another.

[0043] The MFCCs are commonly derived as follows: (1) Take the Fourier transform of a windowed excerpt of a signal. (2) Apply the mel filterbank to the power spectru obtained in (1), sum the energy in each filter. (The mel-scale filterbank is commonly implemented as triangular overlapping windows.) (3) Take the logarithm of all filterbank energies. (4) Take the discrete cosine transform (DCT) of the list of values obtained in (3) to arrive at the MFCCs. The number of the filters in the mel-scale filter bank dictates the number of MFCCs.

[0044] The Delta MFCCs are computed based on the MFCCs as follows:

[0045] To calculate the delta coefficients, the following formula can be used:

$$d_t = \frac{\sum_{n=1}^{N} n(c_{t+n} - c_{t-n})}{2\sum_{n=1}^{N} n^2}$$

where d_t is a delta coefficient, from frame t computed in terms of the MFC coefficients ranging from c_{t+N} to c_{t-N}. A typical value for N is 1 or 2. The number of Delta MFCC is determined by the number of MFCCs.

A person of ordinary skill in the art of speech processing can implement the extraction of formant frequencies and Delta MFCC from a subject’s speech using well-known algorithms described, for example in T.F. Quatieri, Discrete-Time Speech Signal Processing: Principles and Practice, Prentice Hall, 2001 (Chapter 5) and D. Mehta, D. Rudoy, and P. Wolfe. Kalman-based autoregressive moving average modeling and inference for formant and antiformanttracking. The Journal of the Acoustical Society of America, 132(3), 1732–1746, 2012. The relevant portions of these publications are incorporated herein by reference.

[0046] Accordingly, in an example embodiment, the present invention is a method of assessing a condition in a subject. The method comprises measuring at least one speech-related variable in a subject; extracting a channel-delay correlation structure of the at least one speech-related variable; and generating an assessment of a condition of the subject, based on the correlation structure of the at least one speech-related variable. For example, the speech-related variables can include a formant frequency or, for example, at least two formant frequencies. Alternatively or additionally, the at least one speech-related variable can include a Mel Frequency Cepstral Coefficient (MFCC), or a Delta Mel Frequency Cepstral Coefficient (Delta MFCC) or, for example, at least two Delta MFCCs.

[0047] In example embodiments, the channel-delay correlation structure includes channel-delay correlation values and/or channel-delay covariance values. The correlation values and the covariance values can be represented by a channel-delay correlation matrix or a channel-delay covariance matrix, respectively.

[0048] In example embodiments, the method of the present invention can be used to generate an assessment of a condition selected from traumatic brain injury, post-traumatic stress disorder, Parkinson’s disease, Aphasia, Dysphonia, Autism, Alzheimer’s disease, Amyotrophic Lateral Sclerosis (ALS), often referred to as Lou Gehrig’s Disease, stroke, sleep disorders, anxiety disorders, multiple sclerosis, cerebral palsy, and major depressive disorder (MDD). In an example embodiment, the condition is MDD.

[0049] In an example embodiment, the present invention is a method of assessing MDD in a subject, comprising measuring the first three formant frequencies in a subject; extracting from the first three formant frequencies a correlation structure that includes a channel-delay correlation matrix or a channel-delay covariance matrix; and generating an assessment of MDD in the subject, based on the correlation structure.

[0050] In another example embodiment, the present invention is a method of assessing MDD in a subject, comprising measuring the first sixteen Delta MFCCs in a subject; extracting from the first sixteen Delta MFCCs a correlation structure that includes a channel-delay correlation matrix or a channel-delay covariance matrix;

09 Oct 2017

2014374349

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A method of assessing a condition in a subject, the method comprising:

processing an input representing a subject's speech resulting in at least one vocal tract representation of the subject;

extracting, by a structure extractor, a channel-delay correlation structure of the at least one vocal tract representation; and

generating an assessment of a condition of the subject, based on the correlation structure of the at least one vocal tract representation.

2. The method of Claim 1, wherein the at least one vocal tract representation includes a formant frequency.

3. The method of Claim 1 or Claim 2, wherein the at least one vocal tract representation includes two or more formant frequencies.

4. The method of Claim 1, wherein the at least one vocal tract representation includes a Mel Frequency Cepstral Coefficient (MFCC).

5. The method of Claim 1, wherein the at least one vocal tract representation includes a Delta Mel Frequency Cepstral Coefficient (Delta MFCC).

6. The method of Claim 5, wherein the at least one vocal tract representation includes two or more Delta MFCC.

7. The method of any one of Claims 1-6, wherein the channel-delay correlation structure includes at least one of channel-delay correlation values and channel-delay covariance values.

8. The method of any one of Claims 1-6, wherein the channel-delay correlation structure includes at least one of a channel-delay correlation matrix and a channel-delay covariance matrix.

9. The method of any one of Claims 1-8, wherein the condition is selected from traumatic brain injury, post-traumatic stress disorder, Parkinson's disease, Aphasia, Dysphonia, Autism, Alzheimer's disease, Amyotrophic Lateral Sclerosis (ALS), often

referred to as Lou Gehrig's Disease, stroke, sleep disorders, anxiety disorders, multiple sclerosis, cerebral palsy, and major depressive disorder (MDD).

10. The method of any one of Claims 1-9, wherein the condition is MDD.

11. The method of any one of Claims 1-10, wherein:

the at least one vocal tract representation comprise the first three formant frequencies;

a channel-delay correlation structure includes a channel-delay correlation matrix or a channel-delay covariance matrix; and

the condition is MDD.

12. The method of any one of Claims 1-10, wherein:

the at least one vocal tract representation comprise the first sixteen Delta MFCC;

a channel-delay correlation structure includes a channel-delay correlation matrix or a channel-delay covariance matrix; and

the condition is MDD.

13. The method of any one of Claims 1-12, wherein the condition is MDD, and wherein generating the assessment of the condition includes generating an estimate of a Beck score of the subject, an estimate of a Hamilton-D score of the subject, or an estimate of a QIDS score of a subject.

14. The method of Claim 13, further including displaying the estimate of the Beck score, the Hamilton-D score or a QIDS score of the subject.

15. A system for assessing a condition in a subject, the system comprising:

a processing unit that processes an input representing a subject's speech resulting in at least one vocal tract representation of the subject;

a channel-delay correlation structure extractor that extracts a correlation structure of the at least one vocal tract representation; and

an assessment generator that generates an assessment of a condition in the subject based on the correlation structure of the at least one vocal tract representation.

16. The system of Claim 15, wherein the vocal tract representation measuring unit measures a formant frequency.

17. The system of any Claim 15 or Claim 16, wherein the vocal tract representation measuring unit measures at least two formant frequencies.

18. The system of Claim 15, wherein the speech-related variable measuring unit measures a Mel Frequency Cepstral Coefficient (MFCC).

19. The system of Claim 15, wherein the speech-related variable measuring unit measures a Delta Mel Frequency Cepstral Coefficient (Delta MFCC).

20. A system for assessing a condition in a subject configured to perform all the steps of any one of Claims 1-14.

speech-related variable in a subject; extracting a channel-delay correlation structure of the at least one speech-related variable; and generating an assessment of a condition of the subject, based on the correlation structure of the at least one speech-related variable.

[0006] In another embodiment, the present invention is a system for assessing a condition in a subject. The system comprises a speech-related variable measuring unit that measures at least one speech-related variable in a subject; a channel-delay correlation structure extractor that extracts a correlation structure of the at least one speech-related variable; and an assessment generator that generates an assessment of a condition in the subject based on the correlation structure of the at least one speech-related variable.

[006A] In another embodiment, the present invention provides a method of assessing a condition in a subject, the method comprising: processing an input representing a subject's speech resulting in at least one vocal tract representation of the subject; extracting, by a structure extractor, a channel-delay correlation structure of the at least one vocal tract representation; and generating an assessment of a condition of the subject, based on the correlation structure of the at least one vocal tract representation.

[006B] In another embodiment, the present invention provides a system for assessing a condition in a subject, the system comprising: a processing unit that processes an input representing a subject's speech resulting in at least one vocal tract representation of the subject; a channel-delay correlation structure extractor that extracts a correlation structure of the at least one vocal tract representation; and an assessment generator that generates an assessment of a condition in the subject based on the correlation structure of the at least one vocal tract representation.

[006C] Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

[006D] The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

[0007] The methods and the systems described herein are advantageously language-independent. Additional advantages include channel-independence as the methods and systems disclosed herein employ data features that do not change with noise or power.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[0009] The foregoing will be apparent from the following more particular description of example embodiments of the invention, as illustrated in the accompanying drawings in which like reference characters refer to the same parts throughout the different views. The drawings are not necessarily to scale, emphasis instead being placed upon illustrating embodiments of the present invention.

[0010] Figure 1 is a color-coded two-dimensional plot showing an example of a channel-delay correlation matrix computed from formant tracks from a healthy subject (top panel) and from a severely depressed subject (bottom panel).

[0011] Figure 2 is a plot of eigenvalues as a function of eignevalue rank, with eigenvalues ordered from largest to smallest (*i.e.* an eigenspectra), derived from formant channel-delay matrices shown in Figure 1.

DETAILED DESCRIPTION OF THE INVENTION

[0019] A description of example embodiments of the invention follows.

[0020] As used herein, the term "speech-related variable" means an anatomical or a physiological characteristic of a subject that can be measured during the subject's speech and can serve as a basis for generating an assessment of a condition of the subject, as described herein. Examples of speech-related variables include formant frequencies, as defined below, Mel Frequency Cepstral Coefficients (MFCC) and Delta Mel Frequency Cepstral Coefficients (Delta MFCC), as defined below, prosodic characteristics of speech (that is any characteristic of speech that provides information about the timing, intonation, and/or energy), facial features of the speaker, and skin conductance of the speaker.

[0021] Additional examples of speech-related variables include pitch, aspiration, rhythm, tremor, jitter, shimmer, other amplitude- and frequency-modulation functions, as well as their frequency decompositions.

[0022] In some embodiments, certain speech-related variables are referred to herein as "low-level features." Such low-level features include the following. **Harmonics-to-noise ratio (HNR):** HNR is an estimate of the harmonic component divided by the aspiration component in voiced speech, and can act as a measure of "breathiness" in a voice. It is computed over successive frames (e.g., every 10 ms). Aspiration occurs when turbulence is generated at the vibrating vocal folds.

[0023] **Cepstral peak prominence (CPP):** CPP is defined as the difference, in dB, between the magnitude of the highest peak and the noise floor in the power cepstrum for a time interval of greater than about 2 ms and is computed over successive frames (e.g., every 10 ms). (The cepstrum is defined as the Fourier transform of the log-spectrum.) Several studies have reported strong correlations between CPP and overall dysphonia perception, breathiness, and vocal fold kinematics. **Facial action units (FAUs). FAU represent** measurable differences between facial expressions, and relate to facial features derived from optical video of the face that correspond to muscle movements of the face. The facial action coding system (FACS) quantifies localized changes in facial expression representing facial action units (FAUs) that correspond to distinct muscle movements of the face.

[0042] As used herein, the term "Mel Frequency Cepstral Coefficients" (MFCC) refers to the coefficients that collectively make up a "mel-frequency cepstrum" (MFC), which is a representation of the short-term power spectrum of a sound signal. The term "cepstrum" refers to the result of taking the Inverse Fourier transform (IFT) of the logarithm of the spectrum of a signal. The term "mel" refers to the use of the "mel scale" or similar filterbank by the methods that obtain MFCC. The "mel scale" is a perceptual scale of pitches judged by listeners to be equal in distance from one another.

[0043] The MFCCs are commonly derived as follows: (1) Take the Fourier transform of a windowed excerpt of a signal. (2) Apply the mel filterbank to the power spectru obtained in (1), sum the energy in each filter. (The mel-scale filterbank is commonly implemented as triangular overlapping windows.) (3) Take the logarithm of all filterbank energies. (4) Take the discrete cosine transform (DCT) of the list of values obtained in (3) to arrive at the MFCCs. The number of the filters in the mel-scale filter bank dictates the number of MFCCs.

[0044] The Delta MFCCs are computed based on the MFCCs as follows:

[0045] To calculate the delta coefficients, the following formula can be used:

$$d_t = \frac{\sum_{n=1}^{N} n(c_{t+n} - c_{t-n})}{2\sum_{n=1}^{N} n^2}$$

where d_t is a delta coefficient, from frame t computed in terms of the MFC coefficients ranging from c_{t+N} to c_{t-N}. A typical value for N is 1 or 2. The number of Delta MFCC is determined by the number of MFCCs.

A person of ordinary skill in the art of speech processing can implement the extraction of formant frequencies and Delta MFCC from a subject's speech using well-known algorithms described, for example in T.F. Quatieri, Discrete-Time Speech Signal Processing: Principles and Practice, Prentice Hall, 2001 (Chapter 5) and D. Mehta, D. Rudoy, and P. Wolfe. Kalman-based autoregressive moving average modeling and inference for formant and antiformanttracking. The Journal of the Acoustical Society of America, 132(3), 1732–1746, 2012. The relevant portions of these publications are incorporated herein by reference.

09 Oct 2017

2014374349

[0046] Accordingly, in an example embodiment, the present invention is a method of assessing a condition in a subject. The method comprises measuring at least one speech-related variable in a subject; extracting a channel-delay correlation structure of the at least one speech-related variable; and generating an assessment of a condition of the subject, based on the correlation structure of the at least one speech-related variable. For example, the speech-related variables can include a formant frequency or, for example, at least two formant frequencies. Alternatively or additionally, the at least one speech-related variable can include a Mel Frequency Cepstral Coefficient (MFCC), or a Delta Mel Frequency Cepstral Coefficient (Delta MFCC) or, for example, at least two Delta MFCCs.

[0047] In example embodiments, the channel-delay correlation structure includes channel-delay correlation values and/or channel-delay covariance values. The correlation values and the covariance values can be represented by a channel-delay correlation matrix or a channel-delay covariance matrix, respectively.

[0048] In example embodiments, the method of the present invention can be used to generate an assessment of a condition selected from traumatic brain injury, post-traumatic stress disorder, Parkinson’s disease, Aphasia, Dysphonia, Autism, Alzheimer’s disease, Amyotrophic Lateral Sclerosis (ALS), often referred to as Lou Gehrig’s Disease, stroke, sleep disorders, anxiety disorders, multiple sclerosis, cerebral palsy, and major depressive disorder (MDD). In an example embodiment, the condition is MDD.

[0049] In an example embodiment, the present invention is a method of assessing MDD in a subject, comprising measuring the first three formant frequencies in a subject; extracting from the first three formant frequencies a correlation structure that includes a channel-delay correlation matrix or a channel-delay covariance matrix; and generating an assessment of MDD in the subject, based on the correlation structure.

[0050] In another example embodiment, the present invention is a method of assessing MDD in a subject, comprising measuring the first sixteen Delta MFCCs in a subject; extracting from the first sixteen Delta MFCCs a correlation structure that includes a channel-delay correlation matrix or a channel-delay covariance matrix;

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A method of assessing a condition in a subject, the method comprising:

processing an input representing a subject's speech resulting in ~~measuring~~ at least one vocal tract representation of the ~~speech-related variable in a~~ subject;

extracting, by a structure extractor, a channel-delay correlation structure of the at least one vocal tract representation~~speech-related variable~~; and

generating an assessment of a condition of the subject, based on the correlation structure of the at least one vocal tract representation~~speech-related variable~~.

2. The method of Claim 1, wherein the at least one vocal tract representation ~~speech-related variable~~ includes a formant frequency.

3. The method of ~~any one of~~ Claim~~s~~ 1 or Claim 2, wherein the at least one vocal tract representation includes ~~speech-related variables are~~ two or more formant frequencies.

~~4. The method of Claim 1, wherein the at least one speech-related variable includes a facial action unit.~~

4~~5~~. The method of Claim 1, wherein the at least one vocal tract representation ~~speech-related variable~~ includes a Mel Frequency Cepstral Coefficient (MFCC).

5~~6~~. The method of Claim 1, wherein the at least one vocal tract representation ~~speech-related variable~~ includes a Delta Mel Frequency Cepstral Coefficient (Delta MFCC).

6~~7~~. The method of Claim 5~~6~~, wherein the at least one vocal tract representation includes ~~speech-related variables are~~ two or more Delta MFCC.

7~~8~~. The method of any one of Claims 1-6~~7~~, wherein the channel-delay correlation structure includes at least one of channel-delay correlation values and channel-delay covariance values.

~~9. The method of any one of Claims 1-7, wherein the channel-delay correlation structure includes channel-delay covariance values.~~

8~~10~~. The method of any one of Claims 1-6~~9~~, wherein the channel-delay correlation structure includes at least one of a channel-delay correlation matrix and a channel-delay covariance matrix.

2014374349 09 Oct 2017

~~11. The method of any one of Claims 1-9, wherein the channel-delay correlation structure includes a channel-delay covariance matrix.~~

9~~12~~. The method of any one of Claims 1-8~~11~~, wherein the condition is selected from traumatic brain injury, post-traumatic stress disorder, Parkinson's disease, Aphasia, Dysphonia, Autism, Alzheimer's disease, Amyotrophic Lateral Sclerosis (ALS), often referred to as Lou Gehrig's Disease, stroke, sleep disorders, anxiety disorders, multiple sclerosis, cerebral palsy, and major depressive disorder (MDD).

10~~3~~. The method of any one of Claims 1-9~~12~~, wherein the condition is MDD.

11~~4~~. The method of any one of Claims 1-10, wherein:

the at least one vocal tract representation comprise ~~speech-related variables are~~ the first three formant frequencies;

a~~the~~ channel-delay correlation structure includes a channel-delay correlation matrix or a channel-delay covariance matrix; and

the condition is MDD.

12~~5~~. The method of any one of Claims 1-10, wherein:

the at least one vocal tract representation comprise ~~speech-related variables are~~ the first sixteen Delta MFCC;

a~~the~~ channel-delay correlation structure includes a channel-delay correlation matrix or a channel-delay covariance matrix; and

the condition is MDD.

13~~6~~. The method of any one of Claims 1-12~~5~~, wherein the condition is MDD, and wherein generating the assessment of the condition includes generating an estimate of a Beck score of the subject, an estimate of a Hamilton-D score of the subject, or an estimate of a QIDS score of a subject.

14~~7~~. The method of Claim 13~~6~~, further including displaying the estimate of the Beck score, the Hamilton-D score or a QIDS score of the subject.

15~~8~~. A system for assessing a condition in a subject, the system comprising:

a processing ~~speech-related variable measuring~~ unit that processes an input representing a subject's speech resulting in ~~measures~~ at least one vocal tract representation of the ~~speech-related variable in a~~ subject;

a channel-delay correlation structure extractor that extracts a correlation structure of the at least one vocal tract representation~~speech-related variable~~; and

an assessment generator that generates an assessment of a condition in the subject based on the correlation structure of the at least one vocal tract representation~~speech-related variable~~.

16~~9~~. The system of Claim 15~~7~~, wherein the vocal tract representation ~~speech-related variable~~ measuring unit measures a formant frequency.

17~~20~~. The system of any ~~one of~~ Claim~~s~~ 15~~8~~ or Claim 16~~9~~, wherein the vocal tract representation ~~speech-related variable~~ measuring unit measures at least two formant frequencies.

~~21. The system of Claim 17, wherein the at least one speech-related variable includes a facial action unit.~~

18~~22~~. The system of Claim 15~~8~~, wherein the speech-related variable measuring unit measures a Mel Frequency Cepstral Coefficient (MFCC).

19~~23~~. The system of Claim 15~~8~~, wherein the speech-related variable measuring unit measures a Delta Mel Frequency Cepstral Coefficient (Delta MFCC).

20. A system for assessing a condition in a subject configured to perform all the steps of any one of Claims 1-14.

~~24. The system of Claim 23, wherein the speech-related variable measuring unit measures at least two Delta MFCCs.~~

~~25. The system of any one of Claims 18-24, wherein the channel-delay correlation structure extractor extracts channel-delay correlation values.~~

~~26. The system of any one of Claims 1-24, wherein the channel-delay correlation structure extractor extracts channel-delay covariance values.~~

27. The system of any one of Claims 1-24, wherein the channel-delay correlation structure extractor extracts channel-delay correlation matrix.

28. The system of Claim 18, wherein the channel-delay correlation structure extractor extracts a channel-delay covariance matrix.

29. The system of any one of Claims 18-28, wherein the assessment generator generates an assessment of the condition selected from traumatic brain injury, post-traumatic stress disorder, Parkinson's disease, Aphasia, Dysphonia, Autism, Alzeimer's disease, Amyotrophic Lateral Sclerosis (ALS), often referred to as Lou Gehrig's Disease, stroke, sleep disorders, anxiety disorders, multiple sclerosis, cerebral palsy, and major depressive disorder (MDD).

30. The system of any one of Claims 18-29, wherein the assessment generator generates an assessment of MDD.

31. The system of Claim 18, wherein:

the speech-related variable measuring unit measures the first three formant frequencies;

the channel-delay correlation structure extractor extracts channel-delay correlation matrix or a channel-delay covariance matrix; and

the assessment generator generates an assessment of MDD.

32. The system of Claim 18, wherein:

the speech-related variable measuring unit measures the first sixteen Delta MFCC;

the channel-delay correlation structure extractor extracts channel-delay correlation matrix or a channel-delay covariance matrix; and

the assessment generator generates an assessment of MDD.

33. The system of Claim 18, wherein the assessment generator generates an assessment of MDD, and wherein the assessment includes an estimate of a Beck score, a Hamilton-D score or a QIDS score of the subject.

34. The system of Claim 33, further including a display that displays the estimate of the Beck score, the Hamilton-D score, or the QIDS score of the subject.

35. The system of any one of Claims 18-34, wherein the system is a mobile device.