Mobile Health
Lecture 5
Audio Signal and Health

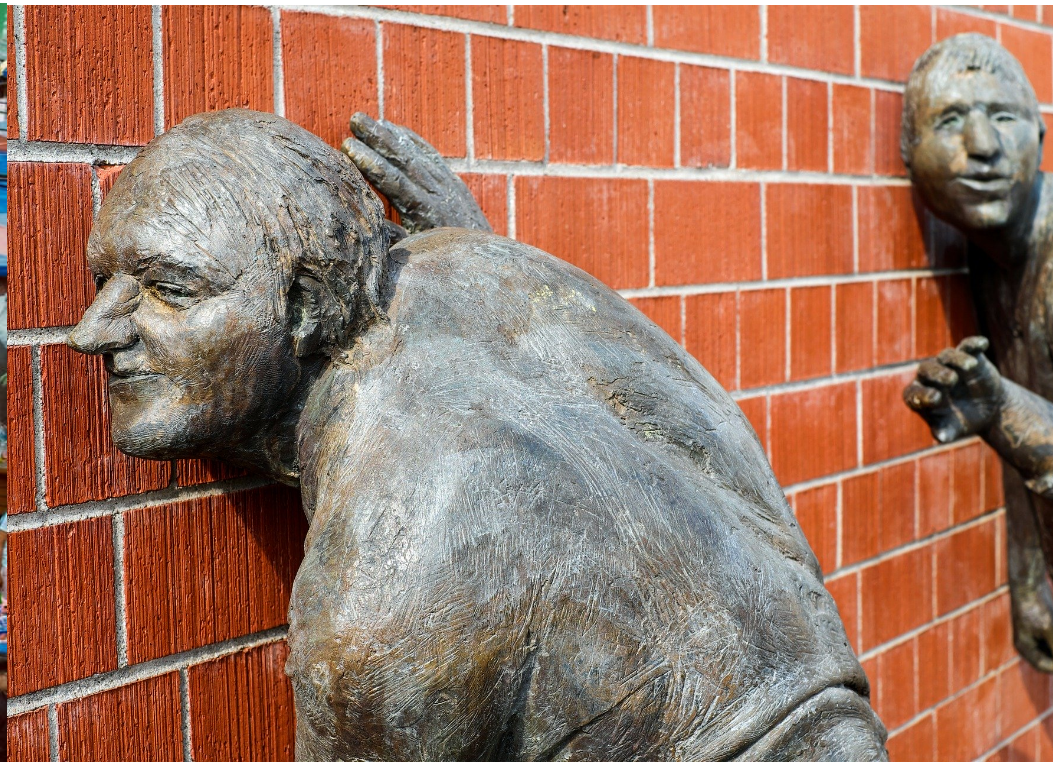
Cecilia Mascolo

UNIVERSITY OF CAMBRIDGE

An Affordable Sensor

UNIVERSITY OF CAMBRIDGE

Automated Sound based Diagnostics



Voice-based Diagnostics

MIT Technology Review

Artificial Intelligence / Machine Learning

Voice Analysis Tech Could Diagnose Disease

Researchers enlist smartphones and machine learning to find vocal patterns that might signal post-traumatic stress disorder or even heart disease.

by **Emily Mullin** Jan 19, 2017

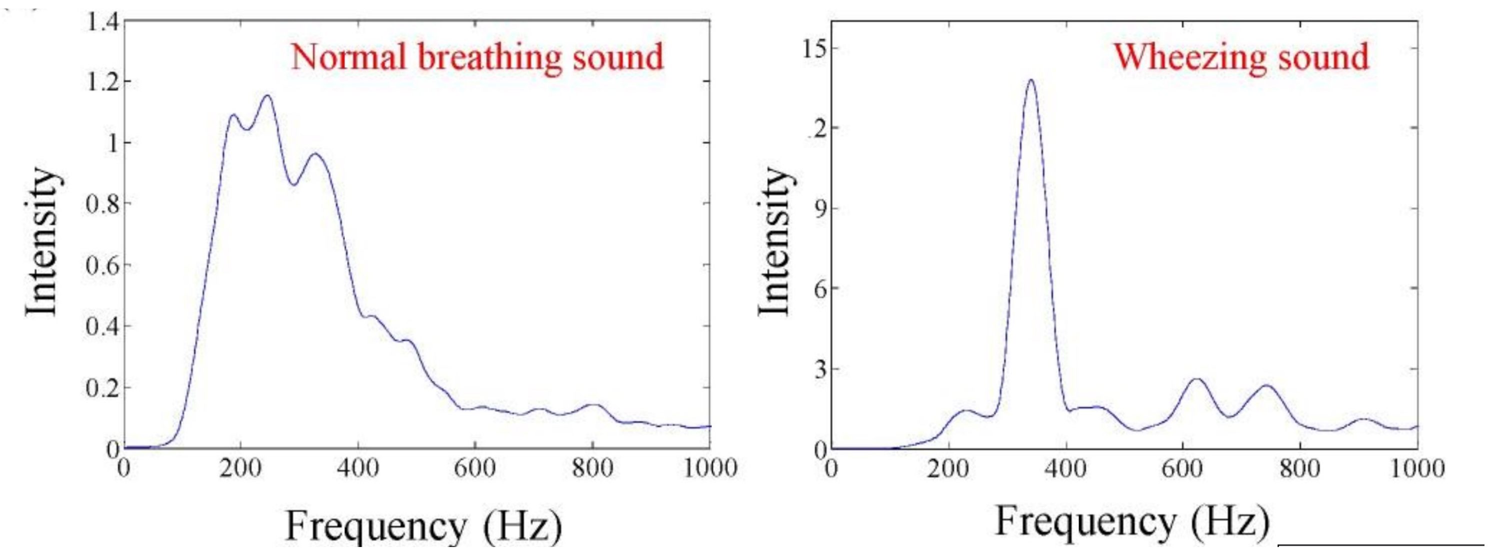
UNIVERSITY OF CAMBRIDGE

Type of diseases for which audio has been tried on

- Depression and PTSD
- Sleep Apnea
- Wheezing (Asthma)
- Parkinson’s
- Alzheimer’s
- Autism
- Cardiovascular: coronary heart disease, arteriosclerosis
- …

Wheezing Detection

- Wheezing sound is a continuously abnormal breathing sound



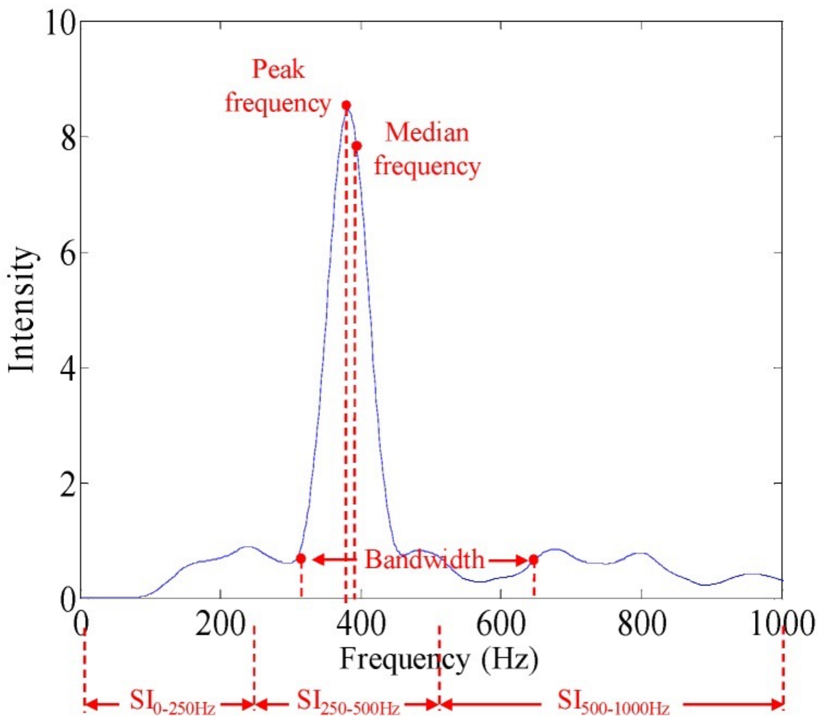
Preprocessing of this signal

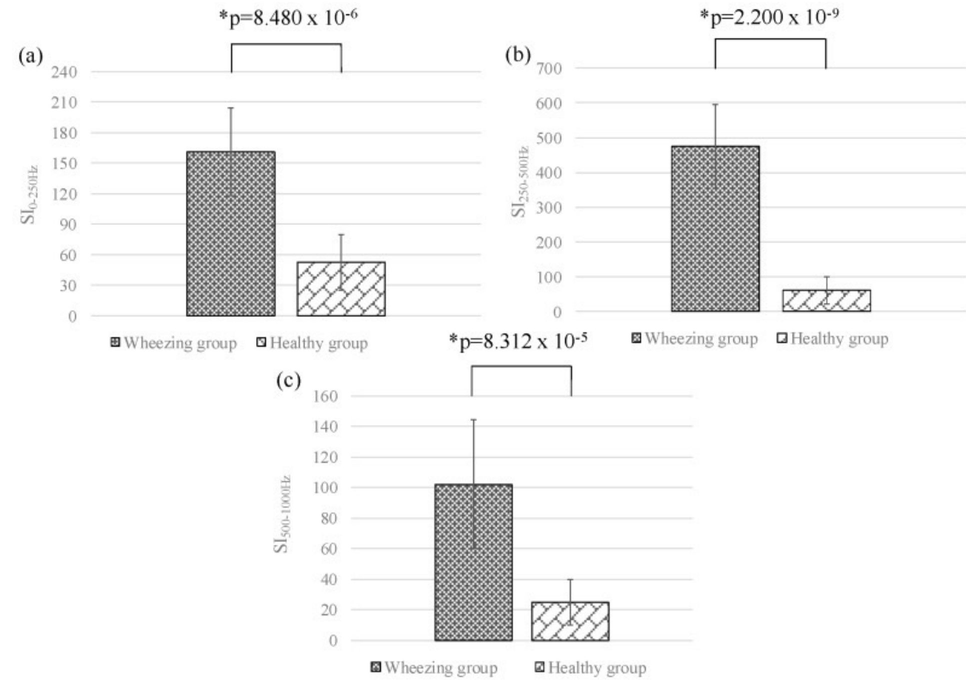
- Band pass filter applied: (150 Hz–1000 Hz)
  - Cut out heart sound, muscle and blood interference sound.
- Raw breathing sound split into 250-ms breathing sound segments with 200-ms overlapping.
- Power spectrums of these breathing sound segments calculated by using DFT with a Hann window.

UNIVERSITY OF CAMBRIDGE

Li SH, Lin BS, Tsai CH, Yang CT, Lin BS. Design of Wearable Breathing Sound Monitoring System for Real-Time Wheeze Detection. Sensors. 2017 Jan 17;17(1):171.

Wheezing Features





UNIVERSITY OF CAMBRIDGE

Li SH, Lin BS, Tsai CH, Yang CT, Lin BS. Design of Wearable Breathing Sound Monitoring System for Real-Time Wheeze Detection. Sensors. 2017 Jan 17;17(1):171.

Important Features

ASE= audio spectral envelope
TI = tonal index
CF1= correlation feature
ER = Energy ratio
K= Kurtosis, Difference to mean ratio
EVD = Eigen Value Decomposition
VC = Vector Comparison
LP = Linear prediction
SPE = Spectral peak entropy
SF = Spectral flatness

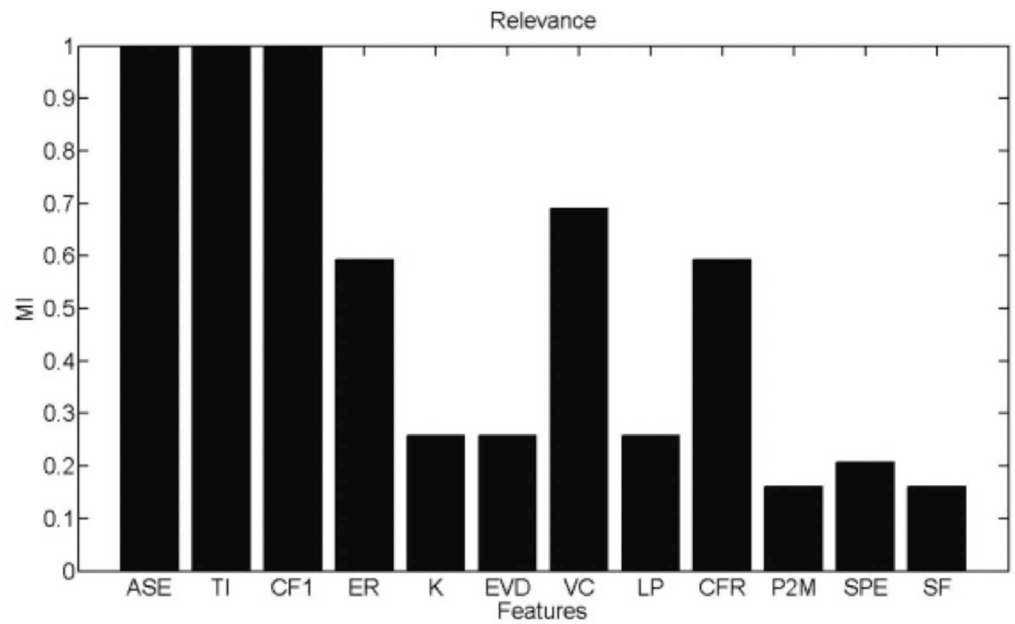


Fig. 2. Relevance of the tested features calculated for real lung recordings using the mRMR algorithm

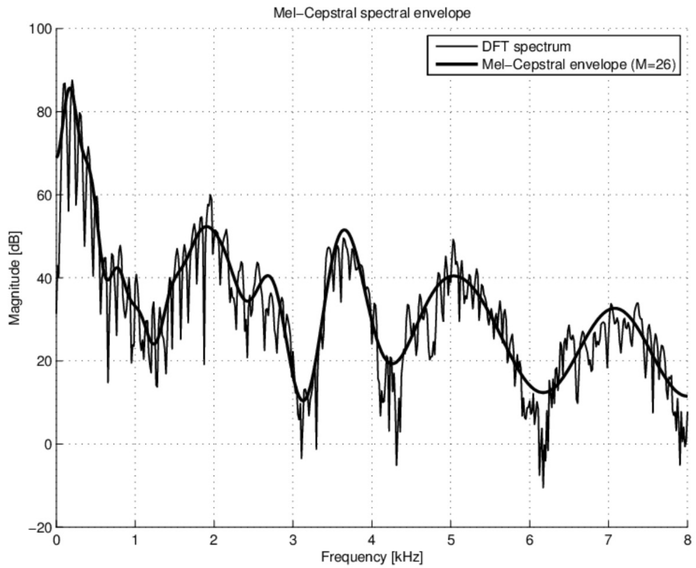
UNIVERSITY OF CAMBRIDGE

MRMR based feature selection for automatic asthma wheezing recognition. M. Wisniewski, T. Zielinski. Signals and Electronic Systems (ICSES), 2012

Frequency Spectrum Envelope

- Identifies the max frequency of each frame and creates a curve which represents all those maximum frequencies.

- Sometimes the frequency spectrum is “smoothened” first.

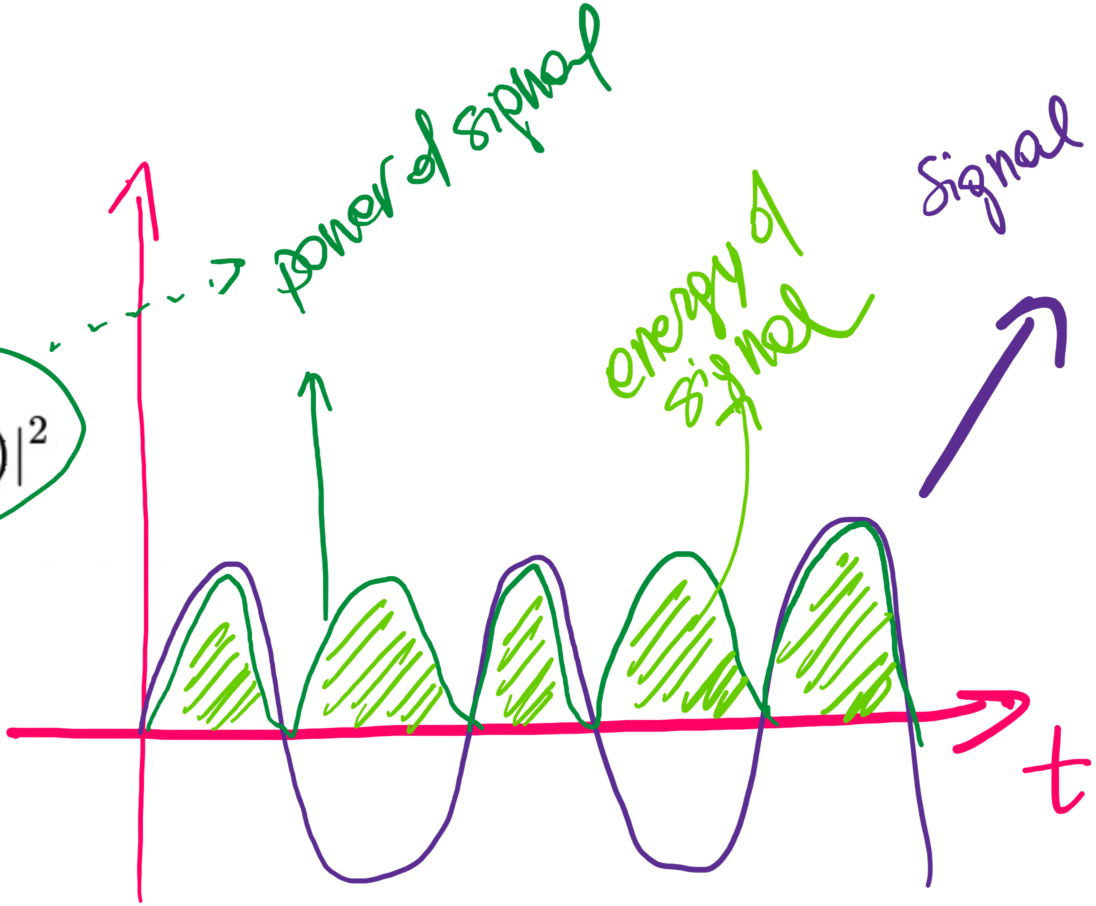


UNIVERSITY OF CAMBRIDGE

Figure from F. Tesser, E. Zovato, P. Cosi (2022). Statistical spectral envelop transformation applied to emotional speech. 13th Int. Conference on Digital Audio Effects. 2010.

Energy of the Signal

$$\sum_{n=-\infty}^{\infty} |x(n)|^2$$



Linear Prediction Coefficients

Technique that calculates coefficients of a linear prediction model which predicts the next sample of audio from a sequence of k previous samples. The coefficient of this linear model are those coefficients. This technique is used to predict **pitch period** accurately (ie the period at which a signal pattern repeats).

MFCCs (Mel-frequency cepstral coefficients) Features

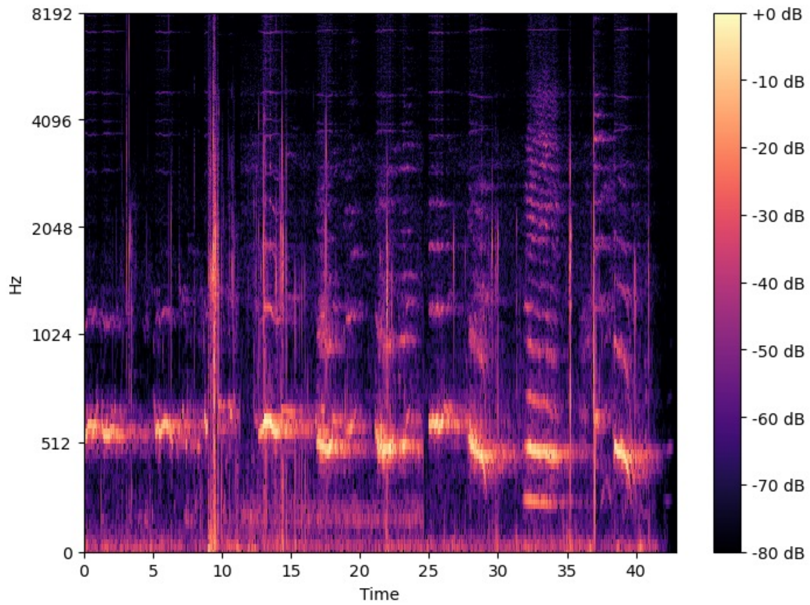
- Very used in audio processing

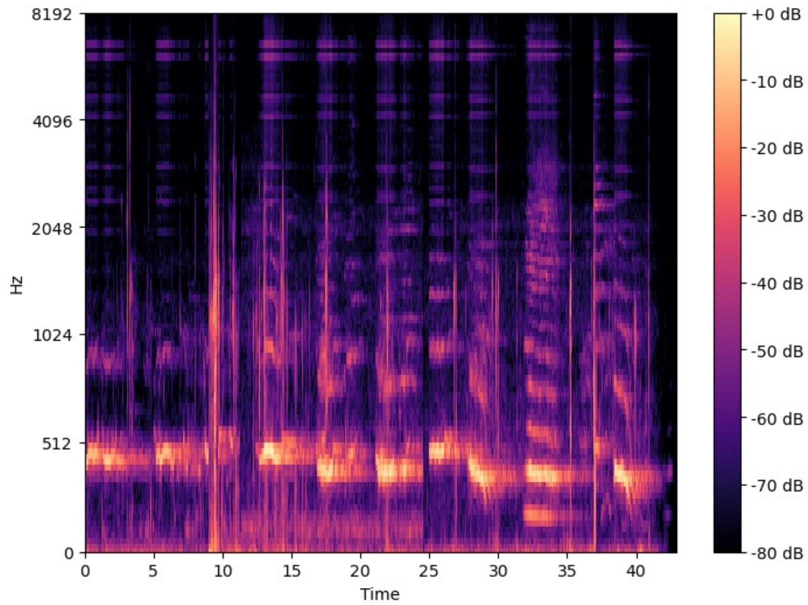
- The intuitively match our auditory way of perceiving sounds

From Spectrograms to Mel Spectrogram

- Spectrogram 1 is the output of standard DFTs.
- However humans perceive frequency of sound "logarithmically"
  - Difference between sounds at lower frequency seems more than difference of sounds at higher frequency

Whale song ☺



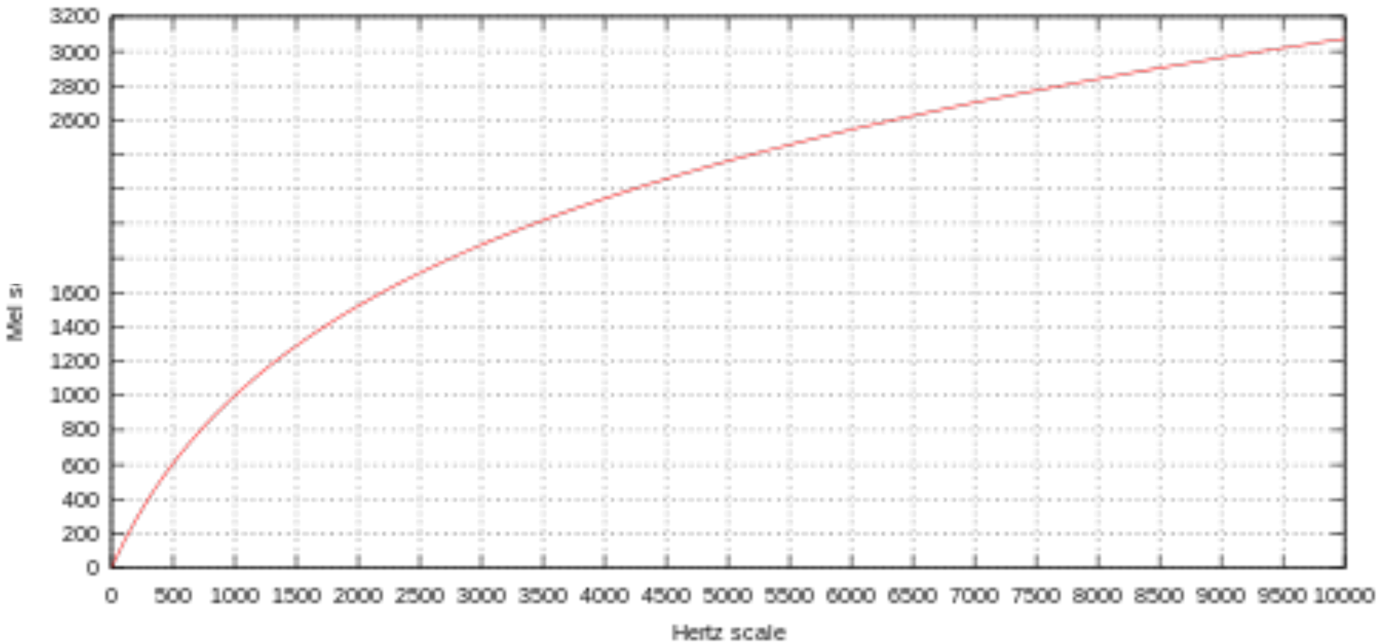


UNIVERSITY OF CAMBRIDGE

Mel Scale

- Mel Scale is perceptually relevant scale for frequency
  - Matching our hearing

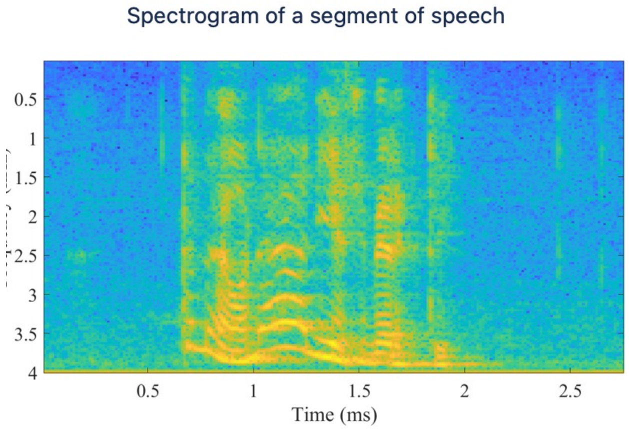
$$m = 2595 \log_{10}\left(1 + \frac{f}{700}\right)$$

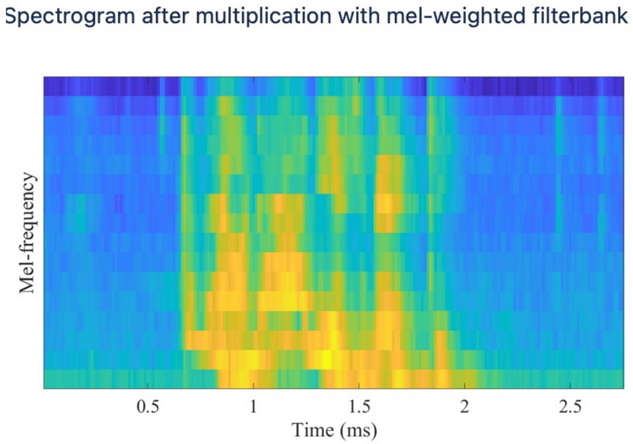


UNIVERSITY OF CAMBRIDGE

Spectrogram to Mel Spectrogram

- Mel spectrograms use Mel Filter banks
  - Effectively bins for the frequency mapping
- Shape retained. Fine structure smoothened.



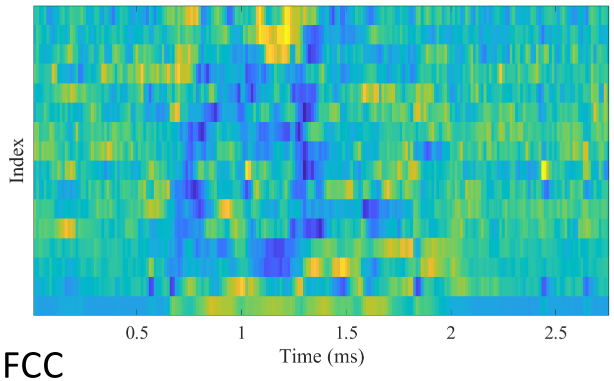


UNIVERSITY OF CAMBRIDGE

Figure from
https://wiki.aalto.fi/display/ITSP/Cepstrum+and+MFCC

Mel Frequency Cepstral Coefficients

- Steps
  - Apply a DFT to generate a frequency plot from time domain plot.
  - Log the data and apply Mel Scaling.
  - Discrete Cosine Transform (a transformation which highlights the important parts of the log DFT graph).
  - Result is a number of "coefficients": first 12-13 are the most relevant generally for audio.

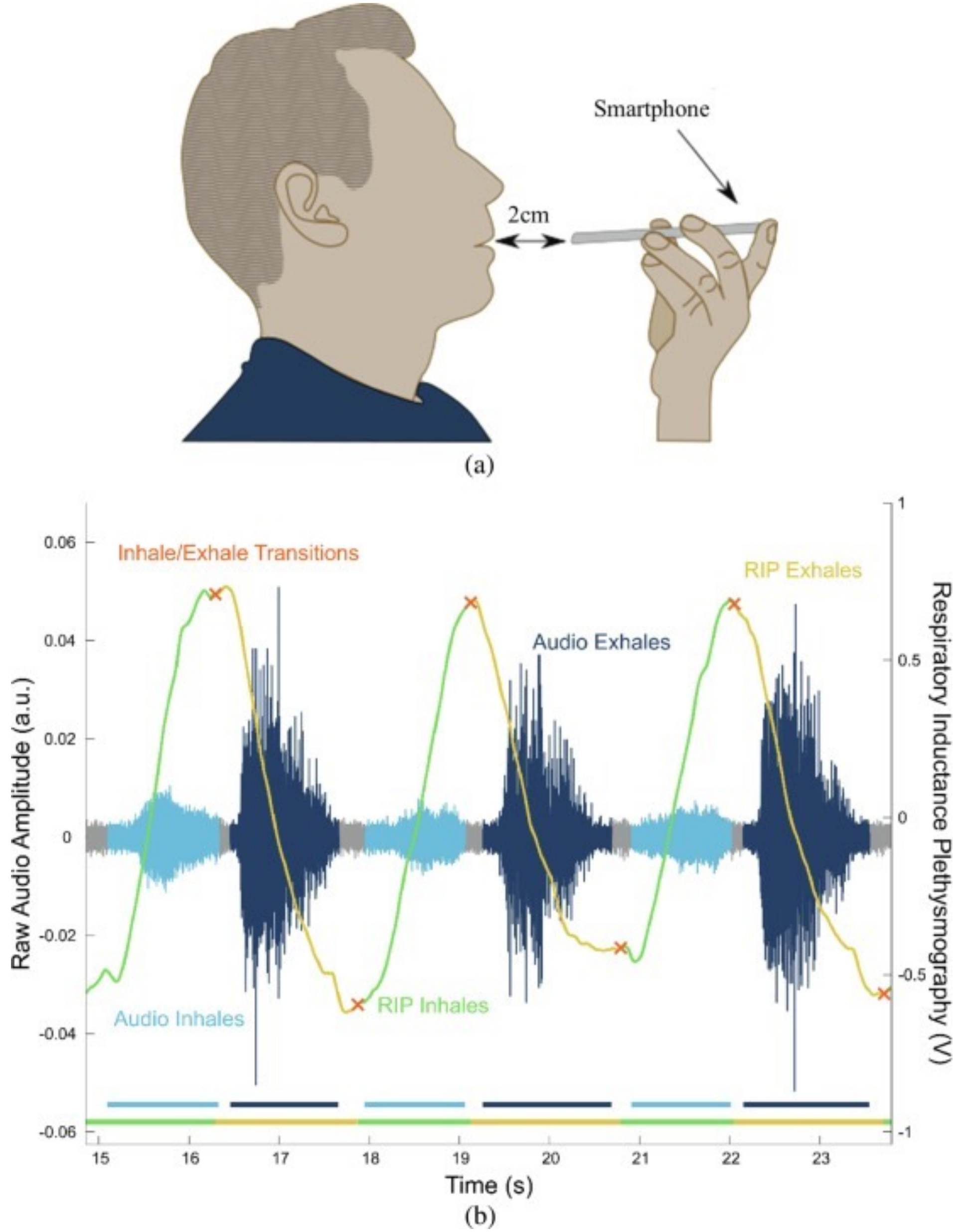


UNIVERSITY OF CAMBRIDGE

Figure from https://wiki.aalto.fi/display/ITSP/Cepstrum+and+MFCC

Respiratory Pattern Detection from Audio

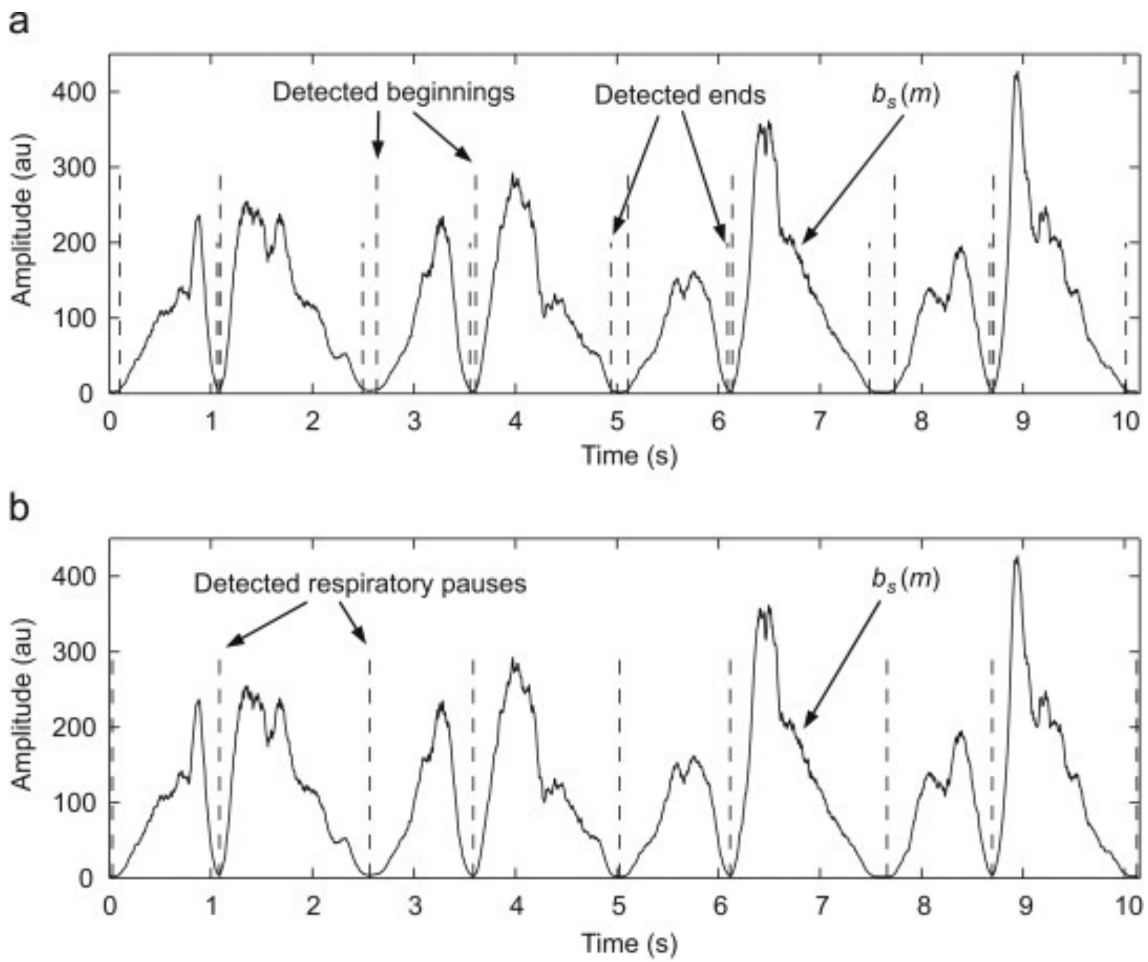
- Green/yellow line is ground truth data from chest strap.
- In-out breathing audio patterns can be visibly distinguished on time domain audio plot.



E Doheny, Ben P.F. O'Callaghan, Vitória S. Fahed, Jérémy Liegey, Cathy Goulding, Silke Ryan, Madeleine M. Lowery, Estimation of respiratory rate and exhale duration using audio signals recorded by smartphone microphones, Biomedical Signal Processing and Control, Volume 80, Part 1, 2023,

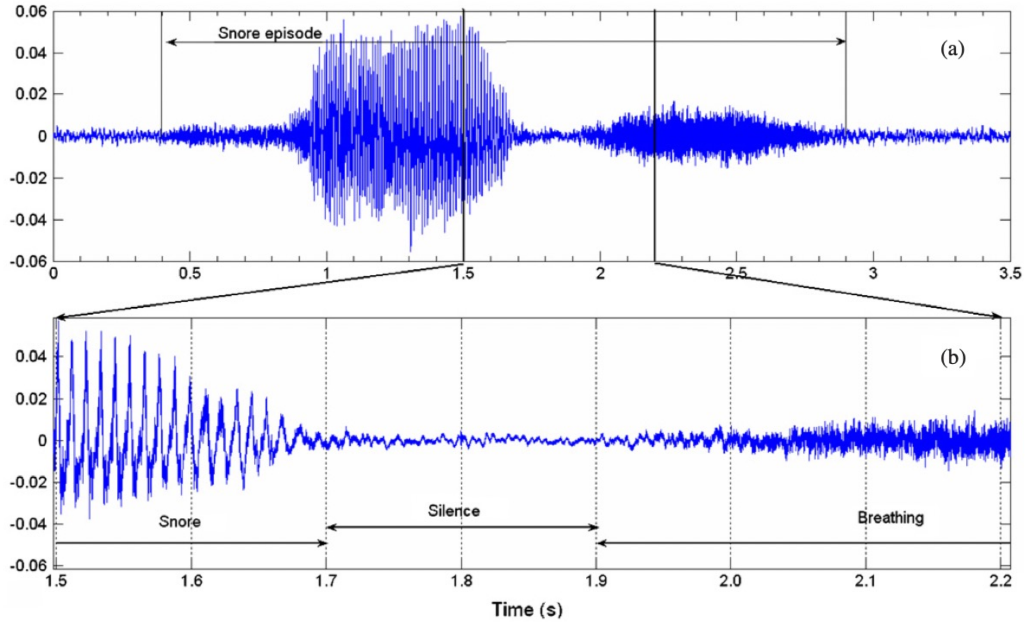
Breathing Pattern Detection

- Signal passed with filters to isolate the right frequencies.
  - 8th order Butterworth low-pass filter with cut-off frequency of 1 kHz.
- Elimination of signal with coughs, yawning etc: windowing and max and median peak amplitude used.
- Time domain features used to detect pauses in respiration.



Snoring-non Snoring

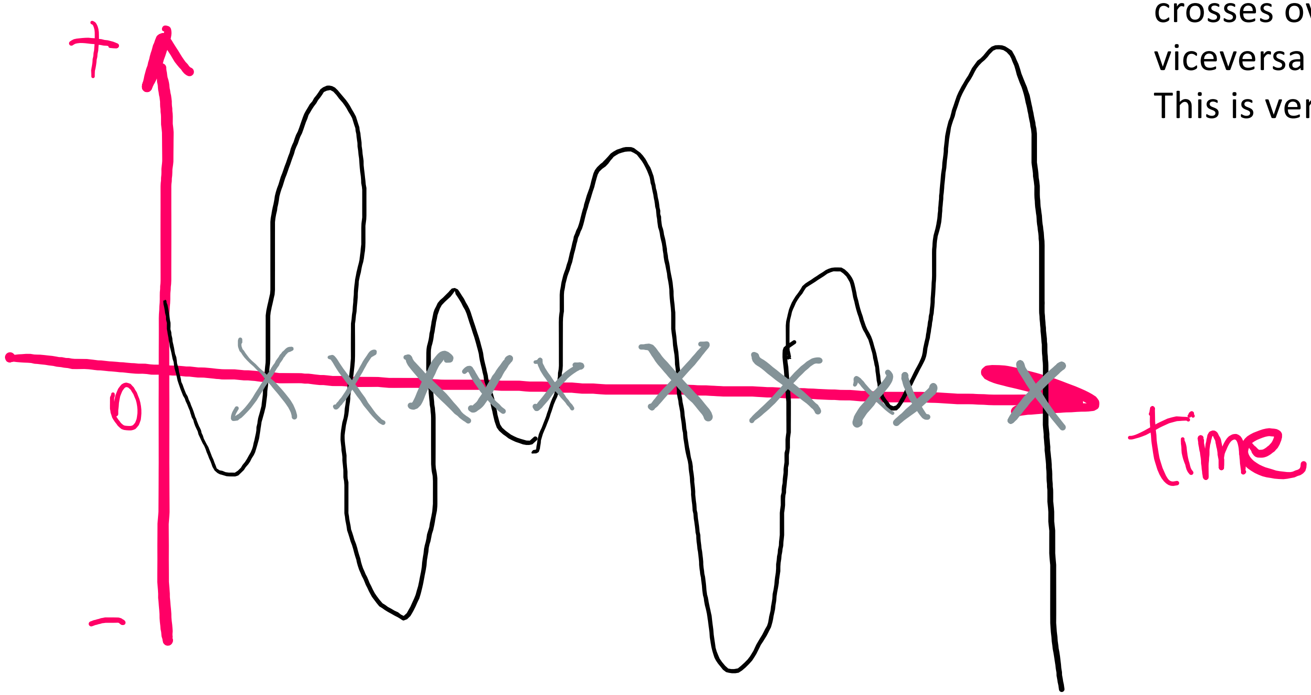
- Identification of voiced snoring, breathing and silence.
- Features
  - number of zero crossings in a given length of time
  - the energy of the signal
  - normalized autocorrelation coefficient at 1 ms delay
  - first predictor coefficient of linear predictive coding (LPC) analysis



UNIVERSITY OF CAMBRIDGE

Karunajeewa AS, Abeyratne UR, Hukins C. Silence-breathing-snore classification from snore-related sounds. Physiol Meas. 2008 Feb;29(2):227-43.

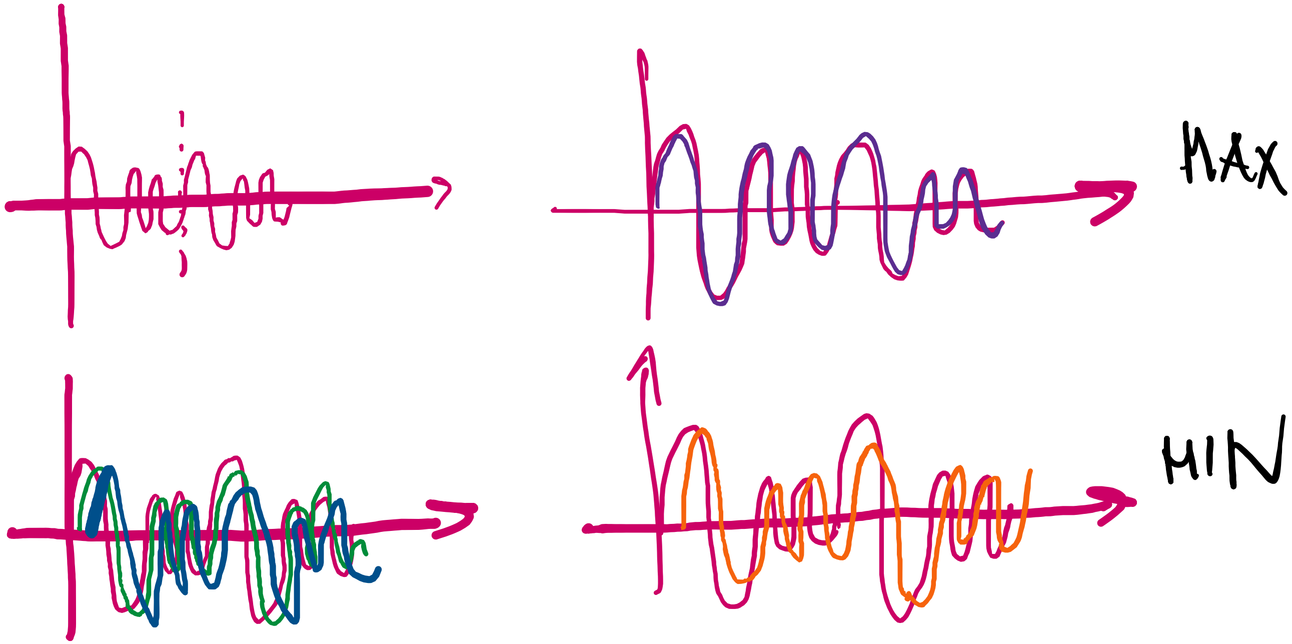
Zero Crossing Rate



Measures the number of times a signal crosses over from positive to negative and viceversa (per time period).
This is very correlated to frequency generally.

Autocorrelation (useful to find period)

$$R(m) = \sum_{n=0}^{N-1-m} s(n)s(n+m)$$

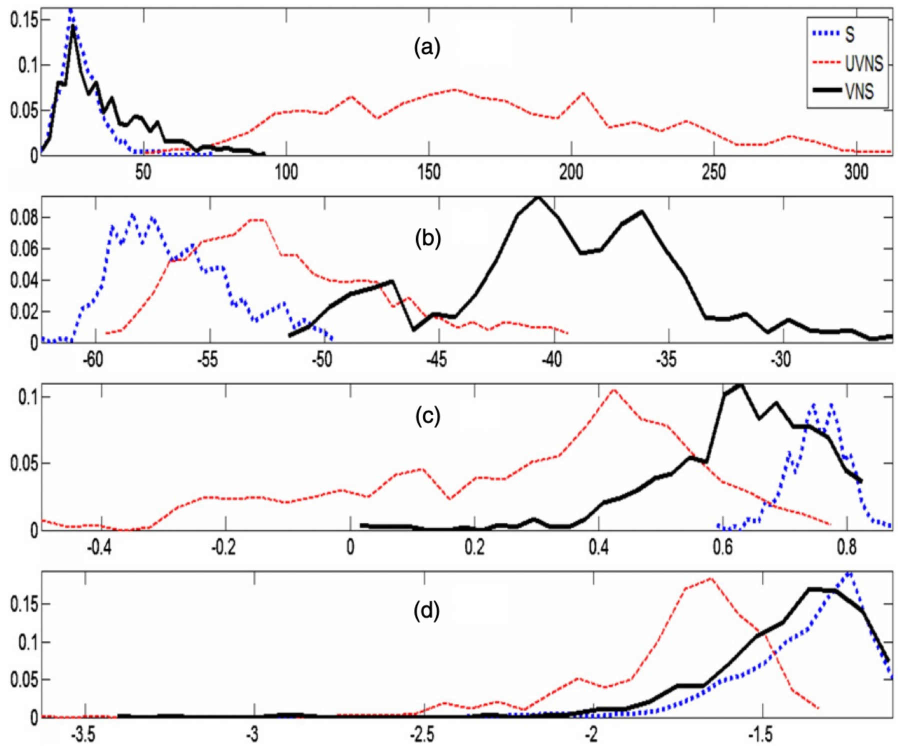


UNIVERSITY OF CAMBRIDGE

Performance: Probability distribution of the features for classification

PDF of
-zero crossing
-log of energy
-autocorrelation coefficient
-first linear predictor coefficient

S= silence
UNS = Unvoiced non silence
VNS = voiced non silence



Sleep Apnea

Cessation of airflow to the lungs that lasts at least for 10s and is associated with at least 4% drop in blood's oxygen saturation level (SaO_2).

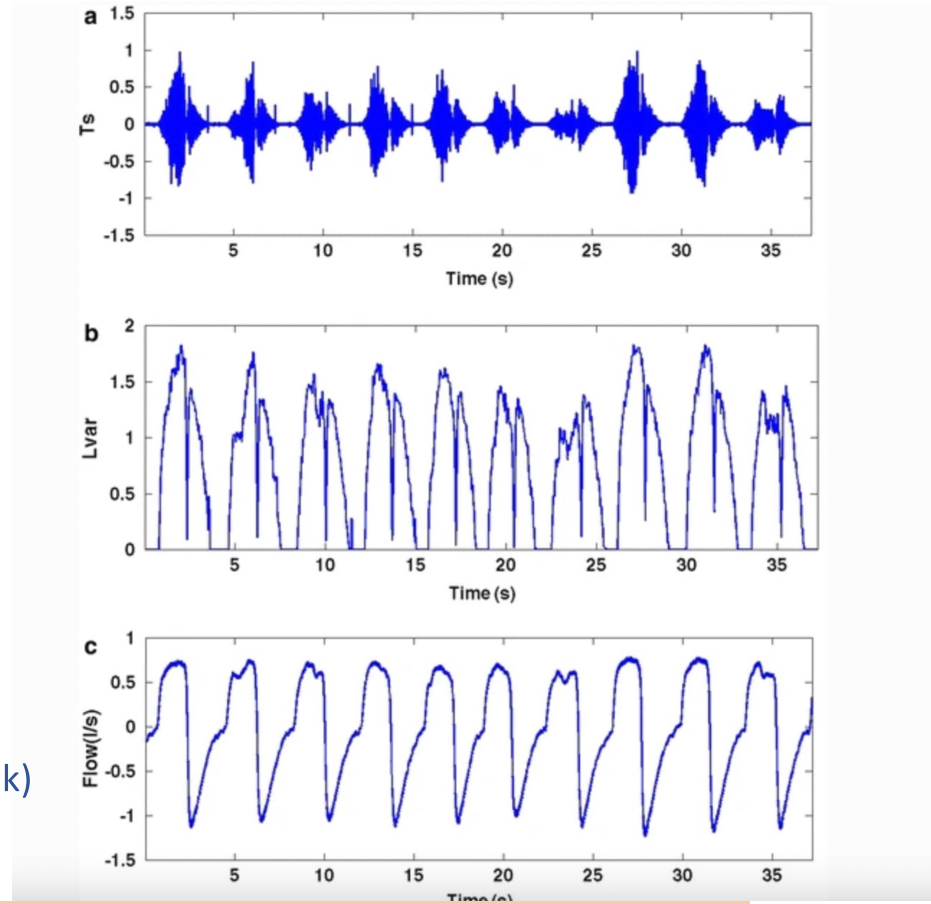
Air Flow and Tracheal Sounds

Log of energy of the tracheal **sound signal** is a good indicator of **air flow**

Signal

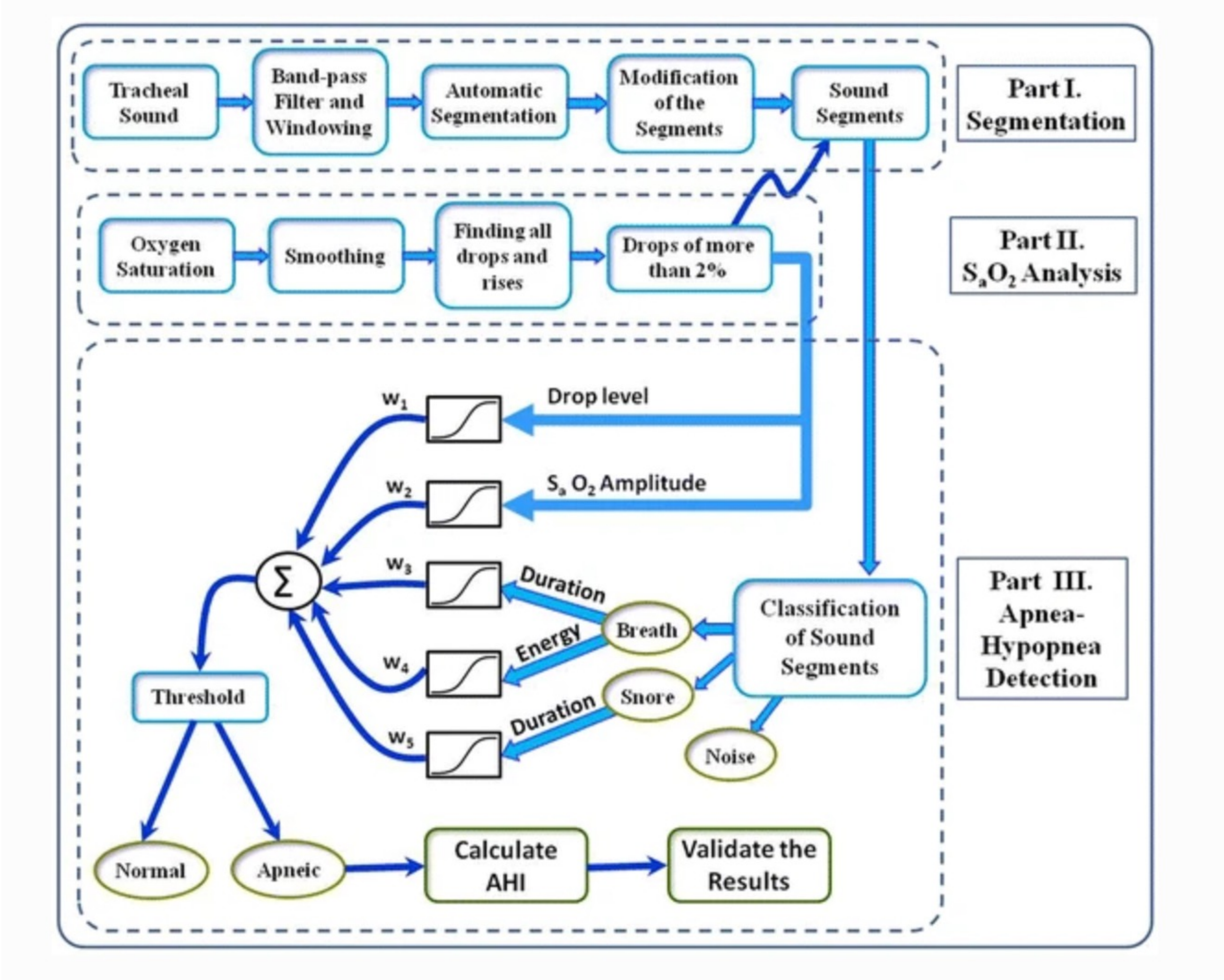
Log energy

True flow
(measured with mask)



Figures from Yadollahi, A., Giannouli, E. & Moussavi, Z. Sleep apnea monitoring and diagnosis based on pulse oximetery and tracheal sound signals. Med Biol Eng Comput 48. (2010).

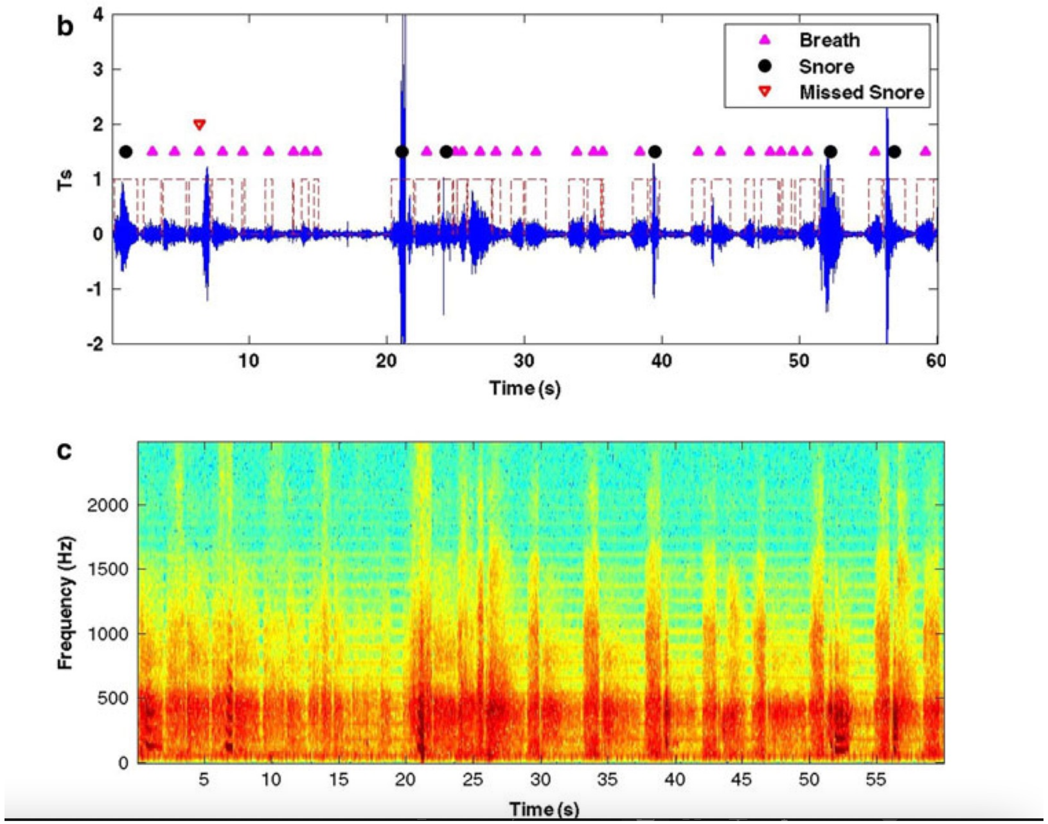
Sleep Apnea Detection with Audio and SaO2 monitoring



UNIVERSITY OF CAMBRIDGE

Sleep Apnea Detection

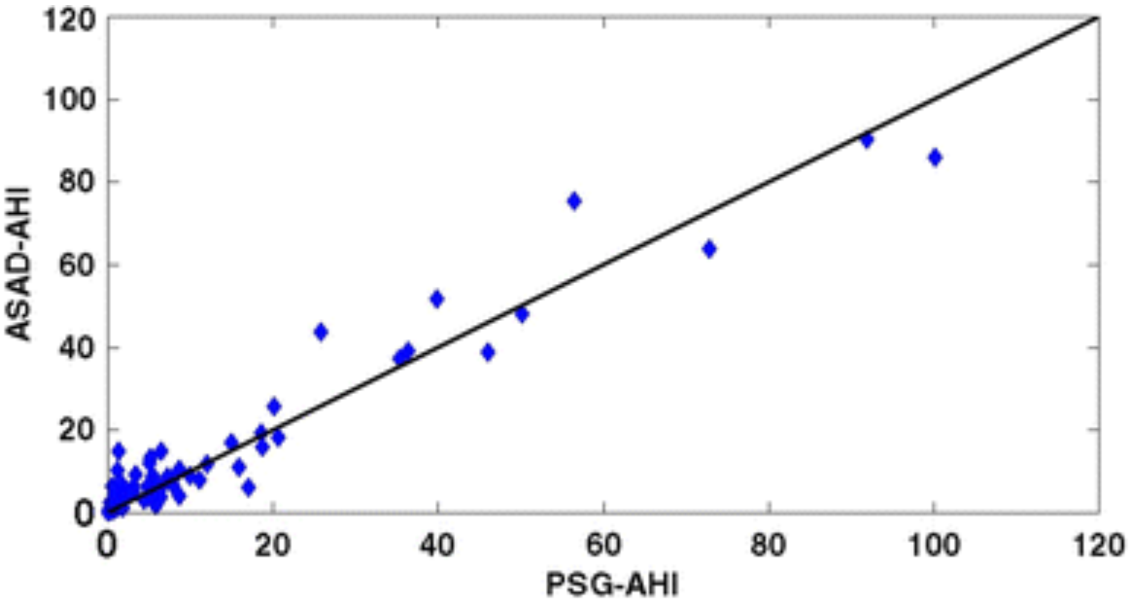
- Classification of the sound in snore, breathing and noise
  - Use of energy and duration (a mixture)
- Spectrogram shows the snoring appearing in deep colours



UNIVERSITY OF CAMBRIDGE

Results

- AHI: apnea–hypopnea index



UNIVERSITY OF CAMBRIDGE

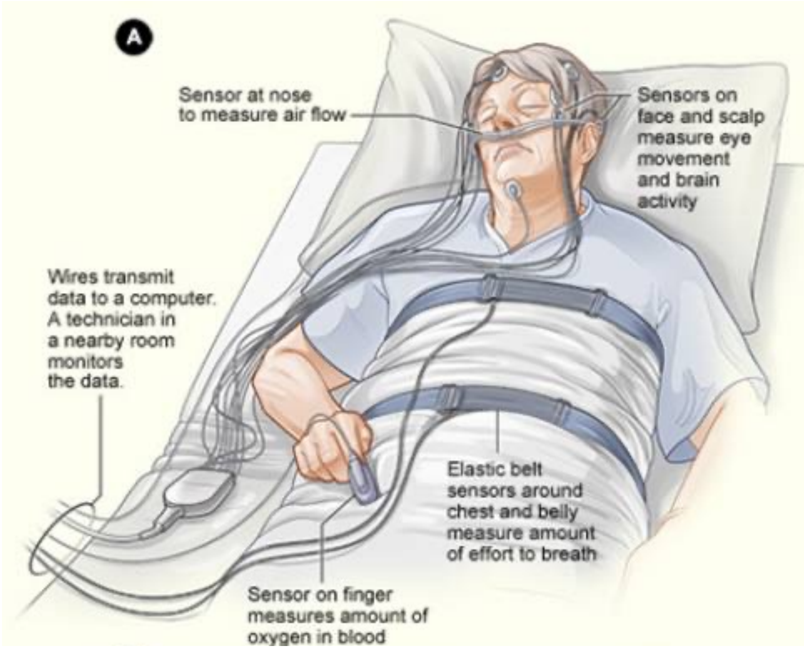
Sleep Stages Classification with Audio

- During sleep (in contrast to wakefulness) there is an increase of upper airway resistance due to decreased activity of the pharyngeal dilator muscles, which is reflected **by amplification of air-pressure oscillations during breathing**. These air-pressure oscillations are perceived as breathing sounds during sleep.
- REM (rapid-eye movement), N(on)REM, and wakefulness are associated with lack of, some, and considerable body movement.
- Breathing pattern is more periodic and consistent in deep NREM sleep compared to REM and wakefulness

Dafna, E., Tarasiuk, A. & Zigel, Y. Sleep staging using nocturnal sound analysis. *Sci Rep* **8**, 13474 (2018).

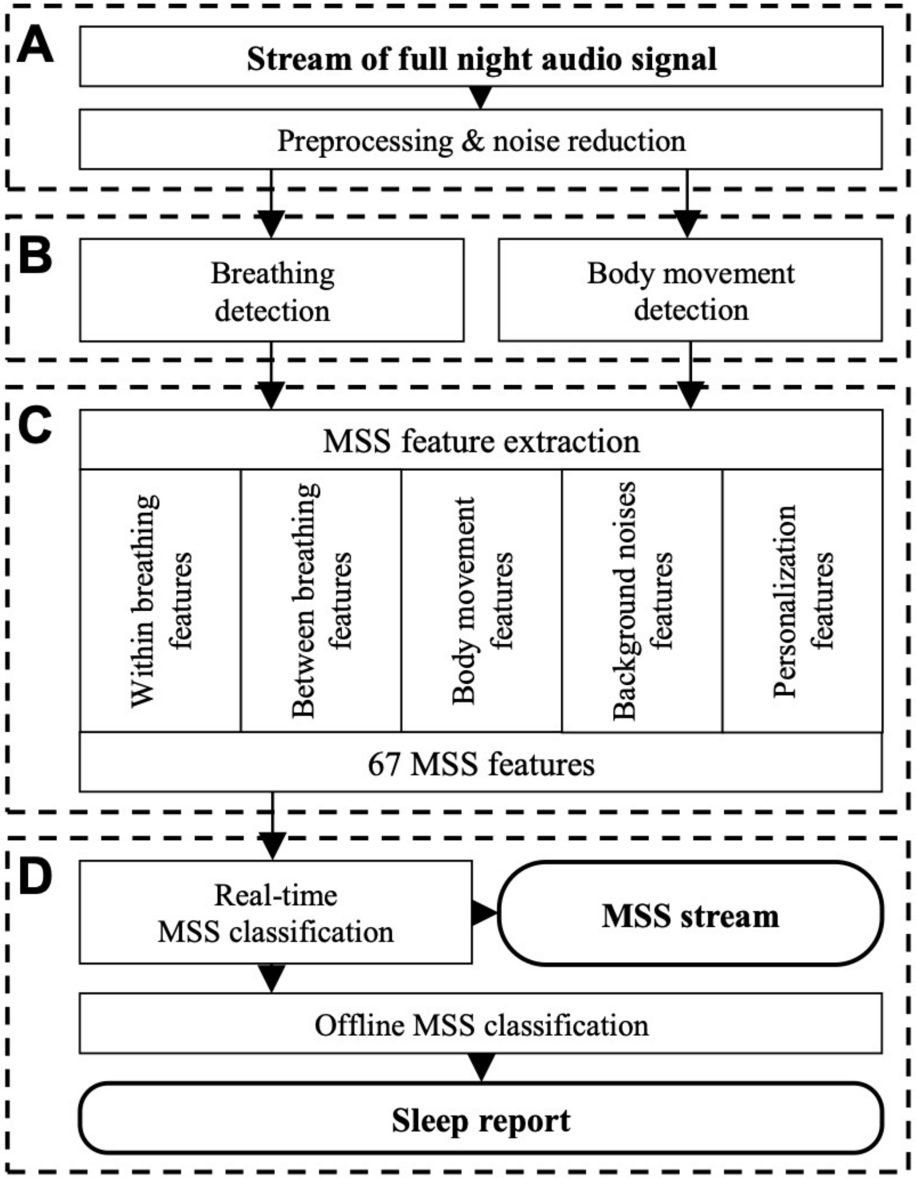
Audio

- Microphone on the bed: (Edirol R-4 pro, Bellingham, WA, USA) with a directional microphone (RØDE, NTG-1, Silverwater, NSW, Australia) was placed at a distance of one meter above the subject's head and used for acquiring the audio signals.
- Polisomnography (PSG) for ground truth



UNIVERSITY OF CAMBRIDGE

Detection of Macro Sleep Stages (MSS)



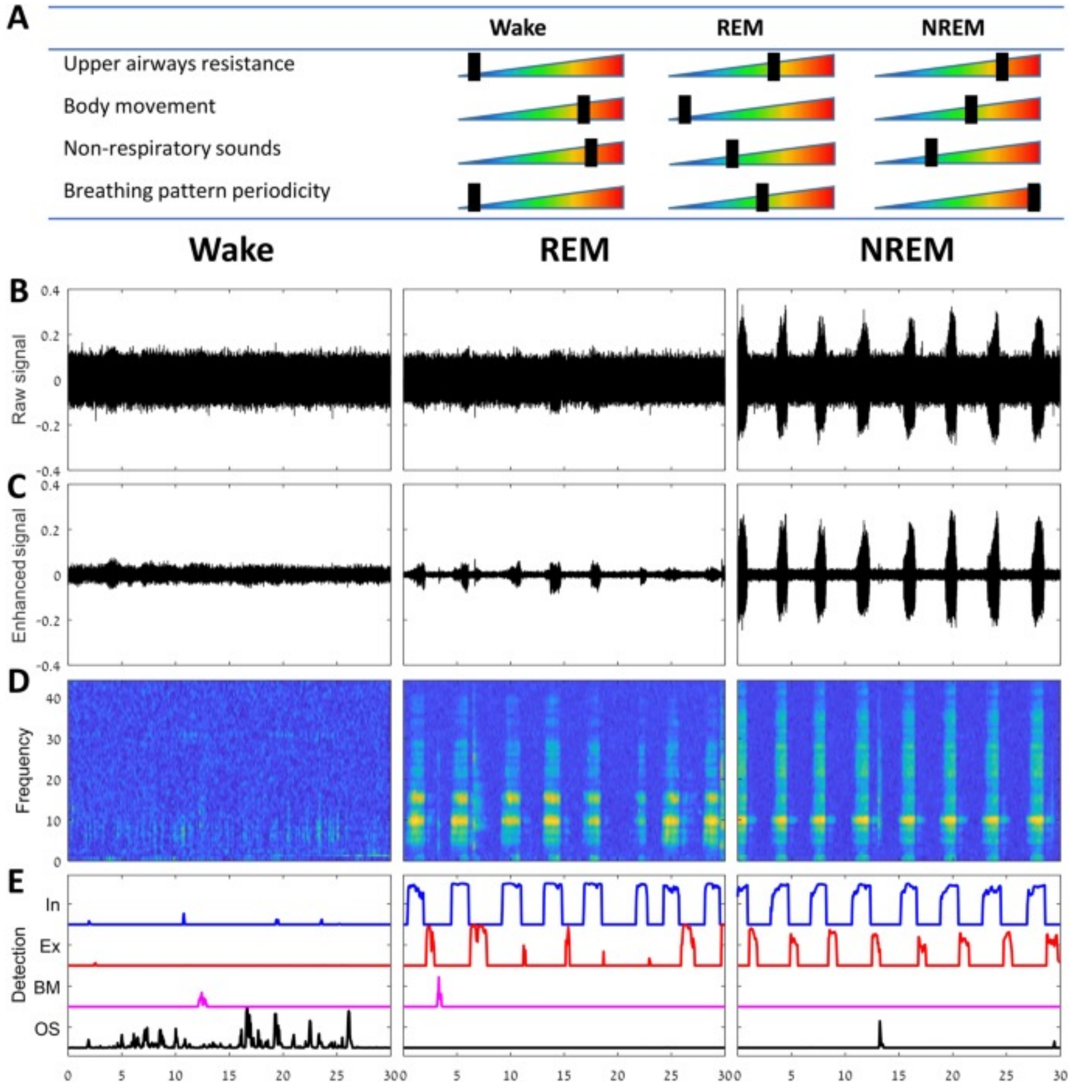
UNIVERSITY OF CAMBRIDGE

Raw sound

Preprocessed

Spectrogram

Inhalation (blue), Exhalation (red), body movement (pink) and other (black)



UNIVERSITY OF CAMBRIDGE

Within Breathing Features

- During sleep, airways resistance is higher than during wakefulness, hence breathing efforts become greater, which translates into several factors including **louder breathing** sounds, prolonged **breathing duration**, and **different vocal sounds** (snores).

		count	importance
A. Within breathing features (WB)	**Feature code**	**33**	**0.270**
Detection score of inspiration (μ,σ)	WB_DI	2	0.093
Detection score of expiration (μ,σ)	WB_DE	2	0.048
Detection score of respiration (μ,σ)	WB_DR	2	0.037
Duration inspiration (μ,σ)	WB_DurI	2	0.075
Duration expiration (μ,σ)	WB_DurE	2	0.024
Stationarity inspiration (μ,σ)	WB_SI	2	0.013
Stationarity expiration (μ,σ)	WB_SE	2	0.009
Sound intensity inspiration (μ,σ)	WB_SII	2	0.044
Sound intensity expiration (μ,σ)	WB_SIE	2	0.009
Sound intensity inspiration top 1% (μ,σ)	WB_SII01	2	0.027
Sound intensity expiration top 1% (μ,σ)	WB_SIE01	2	0.053
Entropy inspiration (μ,σ)	WB_EI	2	0.045
Entropy expiration (μ,σ)	WB_EE	2	0.008
Frequency centroid inspiration (μ,σ)	WB_FCI	2	0.031
Frequency centroid expiration (μ,σ)	WB_FCE	2	0.036
Frequency bandwidth (resp., insp., expi.)	WB_FB	3	0.009

Between Breathing Features

- Alternations in ventilation may affect fundamental respiration factors such as respiratory cycle period, respiratory duty cycle, and respiration consistency, and can be measured using sound analysis. These respiration factors are most likely to have more substantial variability during REM as opposed to NREM.

B. Between breathing features (BB)		**12**	**0.267**
Respiration duty cycle	BB_DCR	1	0.026
Inspiration duty cycle	BB_DCI	1	0.058
Expiration duty cycle	BB_DCE	1	0.020
Respiration cycle period (μ,σ)	BB_RCP	2	0.033
Respiration cycle period consistency	BB_RCPC	1	0.068
Respiration cycle periods fourth-order curve	BB_RCPfit	5	0.023
Breathing Count	BB_BC	1	0.006

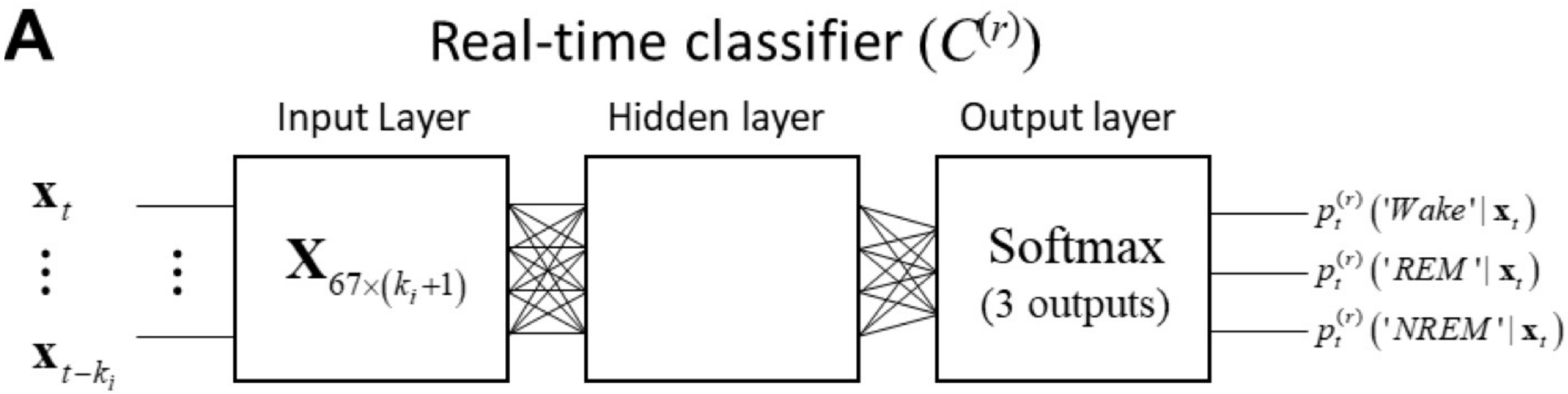
Body Movement Features

- Wakefulness is accompanied by relatively greater body movement, compared to NREM, while during REM sleep body movement should be absent by definition.

C. Body movement features (BM)		**10**	**0.054**
Body movement average score	BM_AS	1	0.002
Body movement overall score percentiles	BM_OS	7	0.017
Sound intensity body movement (all curve)	BM_SI	1	0.007
Sound intensity body movement 10% (all curve)	BM_SI01	1	0.038

UNIVERSITY OF CAMBRIDGE

Real time Classification

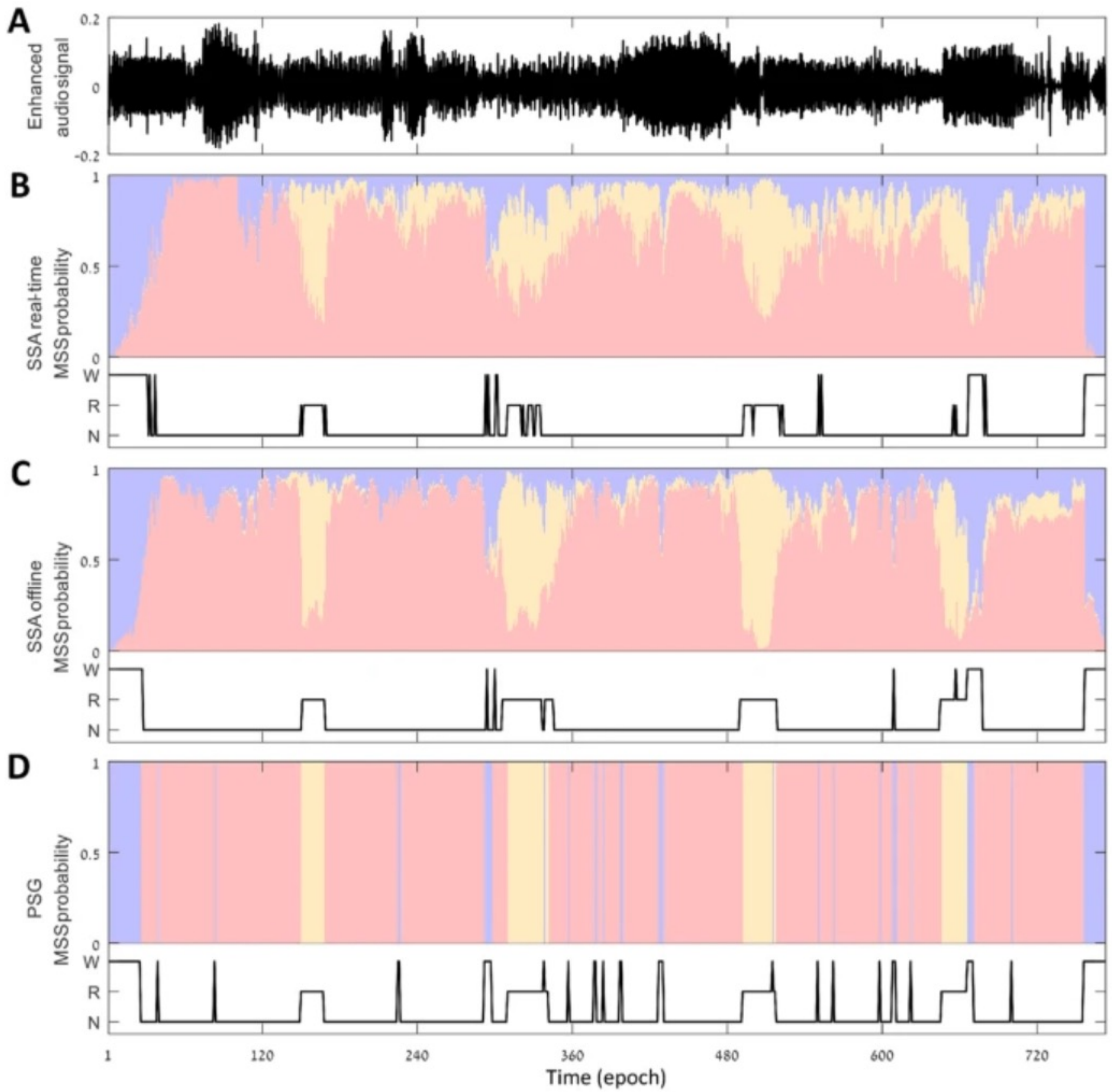


UNIVERSITY OF CAMBRIDGE

Results

One Subject

Blue= wake
Orange= REM sleep
Red= Non-REM sleep



UNIVERSITY OF CAMBRIDGE

Questions

UNIVERSITY OF
CAMBRIDGE