Spectral Feature Extraction and Definitions

William Conley

# The Geneva Minimalistic Acoustic Parameter Set

**Fast-Fourier Transform**

A Fast-Fourier Transform is a process that separates a signal into individual frequencies of different magnitudes. These individual frequencies can also be summed together to re-form the original signal.

**Power Spectrum Density (PSD)**

A Power Spectrum Density graph plots the frequency (Hz) on the x-axis, and power (mV2/Hz) on the y-axis. Using the Fourier transform, the magnitudes (power) of the different frequencies are plotted, showing the distribution of the magnitudes of the individual frequencies that make up signal.

**Long-Term Average Spectrum (LTAS)**

The Long-Term Average spectrum is like the PSD graph, calculated by a fast Fourier transform (NCBI, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5800529/>). In some cases, the long-term average spectrum was calculated using PSD estimation methods, such as Welch’s (<http://ltu.diva-portal.org/smash/get/diva2:1557981/FULLTEXT01.pdf>, p.g. 11).

The difference between an LTAS and a PSD is that the LTAS divides the data into segments, computing an average for each segment, while a PSD considers the data in its entirety.

In the analysis, the PSD graphs are created using the “signal\_psd” function from Neurokit2, which by default uses Welch’s method from SciPy. The SciPy documentation states “Welch’s method computes an estimate of the power spectral density by dividing the data into overlapping segments, computing a modified periodogram for each segment and averaging the periodograms” (<https://docs.scipy.org/doc/scipy-0.14.0/reference/generated/scipy.signal.welch.html>), which lines up with found documentation for LTAS.

All spectral analyses are calculated from an LTAS – “this parameter […] like other spectral slope related parameters – is computed from a logarithmic representation of a band-wise long-term average spectrum (LTAS)” (The Geneva minimalistic acoustic parameter set, p.g. 199)

**Alpha Ratio**

The alpha ratio is a ratio between low-frequency energy and high-frequency energy, as separated by a frequency selected from the distribution to serve as a threshold (The Geneva minimalistic acoustic parameter set, p.g. 199). A PSD graph can be used to calculate this ratio by summing the power of frequencies lower than the threshold and dividing that by the sum of the power of frequencies greater than the threshold. For example, if the threshold was at 80 Hz, the calculation would be:

The article suggests taking a ratio of the frequencies between 50-1000 Hz and 1000-5000 Hz. Since the data we are working with only goes up to 450 Hz, we could choose a different threshold instead.

Since our analysis uses EMG signals rather than voice signals, the frequencies are divided differently. There is no “high frequency” and “low frequency” voice signals that can be divided, but we can instead work with fast-twitching and slow-twitching muscles. Literature suggests that the frequency of slow-twitching muscles is < 60 Hz, and the frequency of fast-twitching muscles is > 60 Hz (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7242282/>), so we will use a threshold of 60 Hz when calculating the alpha ratio, and instead refer to it as the “twitch ratio”.

**Hammarberg Index**

The Hammarberg index is a ratio comparing the highest peak recorded in a low-Hz region – the 0-2000 Hz range, and the highest peak recorded in a high-Hz region – the 2000-5000 Hz range (The Geneva minimalistic acoustic parameter set, p.g. 199). This is like the alpha ratio, however instead of comparing the means, we are comparing the max values. A PSD graph can be used to calculate this in a similar way to the alpha ratio:

Since the data we have does not record frequencies above 450 Hz, we could instead choose a separate threshold to divide “high” and “low” frequencies (perhaps the same one used in the alpha ratio).

Using the same threshold as the “twitch ratio”, we will again use a threshold of 60 Hz, and refer to this calculation as the “twitch index”.

**Spectral Slope**

The spectral slope is calculated by finding the linear least squares approximation of an LTAS (The Geneva Minimalistic Acoustic Parameter Set, p.g. 199).

The spectral slope is calculated for both the lower and higher frequencies, as used in the Alpha Ratio and Hammarberg Index. Similarly, we can calculate the spectral slope for the slow-twitching and fast-twitching muscles.

**Spectral Energy Proportions**

The spectral energy proportions are calculated by dividing the energy in the high-frequency bins and the low-frequency bins by the total amount of energy in the LTAS (The Geneva Minimalistic Acoustic Parameter Set, p.g. 199).

As done previously, this can be adapted muscles by performing the same calculation but using the muscle-twitching threshold as the separation.

**Spectral Flux**

The spectral flux is calculated by dividing the signal into frames, taking a PSD graph of each frame, and summing the squared difference of the PSDs of each consecutive frame (The Geneva Minimalistic Acoustic Parameter Set, p.g. 200).

# EMG Pattern Recognition in the Era of Big Data and Deep Learning

Before each of the following features were extracted, the EMG signals had basic filtering applied (low-pass filters, Butterworth filters, etc.). No full wave rectifiers or smoothing functions were applied. (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881049/>).

**Zero Crossings**

The zero crossings (ZC) are the number of times that a signal crosses 0 over a given sample (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881049/>). To avoid counting zero crossings caused by noise, a threshold is typically included to filter out zero crossings that don’t reach above a specified amplitude. In a study (<https://www.sciencedirect.com/science/article/abs/pii/S1746809418301447>), researchers found that no threshold is needed to calculate the zero crossings and slope sign changes of surface EMG signals, such as the zygomatic and corrugator muscles being studied in this project.

Zero crossings provide a rough estimation of the dominant frequency in the spectrum, and the spectral centroid (<https://www.sciencedirect.com/topics/engineering/zero-crossing-rate>). Provided we have already calculated the LTAS of the signal, we could instead just calculate the maximum of the LTAS, and calculate the spectral centroid separately, as zero crossings are very sensitive to error.

Zero crossings also provide important information about the frequency domain characteristics of the signal, and is an important indicator of muscle fatigue (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8623265/>).

**Spectral Centroid**

The spectral centroid is a measure of the “center of gravity” of the spectrum (<https://www.sciencedirect.com/topics/engineering/spectral-centroid>):

**Spectral Spread**

The spectral spread is the central moment of the spectrum (<https://www.sciencedirect.com/topics/engineering/spectral-centroid>):

**Slope Sign Changes**

Slope sign changes (SSC) are the number of times that a signal changes signs over a given sample (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881049/>). A threshold is also used for this feature to eliminate noise, but as previously stated (<https://www.sciencedirect.com/science/article/abs/pii/S1746809418301447>) the threshold is not needed for surface EMG signals as in this project.

Slope sign changes provide important information about the frequency domain characteristics of a signal (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8623265/>).

**Willison Amplitude**

The Willison Amplitude (WAMP) is the number of times an EMG signal exceeds a threshold, thusly indicating the muscle has been triggered (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881049/>). A study (<https://www.sciencedirect.com/science/article/abs/pii/S1746809418301447>) suggests using a low threshold of R = 0.1:1 times the average root mean square of the baseline.

**Mean Absolute Value**

Mean absolute value (MAV) is the mean absolute value of a signal with N samples, where

Is the kth sample in the window (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881049/>):

**Variance**

Variance is a measure of the EMG signal’s power (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881049/>):

**V-Order**

The v-order is an estimation of the force exerted by the muscle (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881049/>). A study has shown that given the optimal values chosen for the parameters of the v-order, it ends up being the same as the square root of the variance:

**Log-Detector**

The log detector is also an estimate of the force exerted by the muscle (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881049/>). It is defined as:

**EMG Histogram**

The EMG histogram shows how often the EMG signal reaches certain amplitudes. In each signal, the minimum and maximum voltage of the signal is taken and used to make a histogram with 9 bins. This histogram is populated with the amplitudes of each wave in the signal (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881049/>).

The following graph compares classification accuracy of each feature with the amount of disturbance placed on the electrodes (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881049/>):

A graph of different levels of a train

Description automatically generated with medium confidence

The article did not say what the zero crossings and other features were predicting.

# ~~Speech Emotion Recognition – Emotional Models, Databases, Features, Preprocessing Methods, Supporting Modalities and Classifiers~~

**~~Mel Frequency Cepstral Coefficients (MFCC)~~**

~~The MFCC is a descriptor for “timbre”, a feature of audio signals. This may not translate to a useful feature, but we could always include it for added dimensionality (~~[~~https://medium.com/@derutycsl/intuitive-understanding-of-mfccs-836d36a1f779~~](https://medium.com/@derutycsl/intuitive-understanding-of-mfccs-836d36a1f779)~~).~~

~~Perhaps we could have different classes of features. One set we could say we know for sure are features of EMG signals, and another set could be technically possible, but without certainty they have any useful information. That way, people using the dataset could choose to use only the primary features, or investigate the secondary to see whether they hold any value when added to the model.~~

# Surface Electromyography Signal Processing and Classification Techniques

**Justifications**

<https://blog.nasm.org/fitness/fast-twitch-vs-slow-twitch>

This website explains the difference between slow-twitch muscles and fast-twitch muscles. Fast-twitch muscles are quicker and more powerful, but only for a small amount of time. Slow-twitch muscles make smaller movements but can be sustained for much longer. The type of muscles you have depend on how much they are exercised. An average non-athletic person has about 50/50 slow to fast twitch muscles, but this can be changed through different forms of training. Power athletes tend to have more fast-twitch fibers, while endurance athletes tend to have more slow-twitch muscles. While this could cause a discrepancy in the results from different people, we are dealing with facial muscles, which are not typically exercised in the same way for fitness. However, it is worth noting that age is a factor is muscle type as well, with a decline in fast-twitch fibers as a person ages.

<https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/fast-muscle-fiber>

This book suggests that fast-twitch fibers contract at a frequency of at least 60 Hz/s, and slow-twitch fibers contract at a frequency of at least 16 Hz/s. We could use 60 Hz as the threshold, with < 60 being slow-twitching and > 60 being fast-twitching.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7242282/>

This website explores how the activation level of the bicep muscle changes when completely fatigued. It uses EMG software and Fourier analysis to separate the lower and higher frequency components, calculates the intensity of each, and determine the activation level of each category. The trials recorded a range of 5-350 Hz, with the range of 50-60 Hz being filtered out. The study found that as the fatiguing protocol proceeded, the activity of low-frequency muscle fibers rose, while the activity of high-frequency muscle fibers lowered. In the study, 60 Hz appeared to be the cutoff separating high-frequency muscle fibers from low-frequency muscle fibers. This follows the results found in the previous book examined.

<https://eu-ireland-custom-media-prod.s3.eu-west-1.amazonaws.com/UKMEAEU/eSample/9780323755115.pdf>

This website justifies that there are both fast-twitch and slow-twitch muscles in human faces.

<https://medium.com/@davidpinyol91/hammarberg-index-implementation-in-python-660e52a722e>

This website documents using the SciPy Welch method for calculating the Hammarberg Index in Python.

<https://dsp.stackexchange.com/questions/14699/long-term-average-spectrum-of-large-batch-of-audio-files>

This comment suggests using the Welch method for calculating an LTAS.

<https://www.researchgate.net/publication/317098414_Long-term_Average_Spectrum_in_Popular_Music_and_its_Relation_to_the_Level_of_the_Percussion>

This article confirms that the LTAS is calculated using overlapping windows, taking the mean of each (p.g. 3, section 2.2)

<https://www.maths.lu.se/fileadmin/maths/personal_staff/Andreas_Jakobsson/StoicaM05.pdf>

This textbook confirms that the Welch approximation for a PSD is calculated by averaging overlapping windows (p.g. 50, section 2.7.2)

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881049/>

This website gives formulas for mean absolute value (MAV), zero crossings (ZC), slope sign changes (SSC), waveform length (WL), Willison amplitude (WAMP), variance (VAR), v-order, log-detector, EMG histogram, autoregression coefficient (AR), cepstrum coefficient (CC),

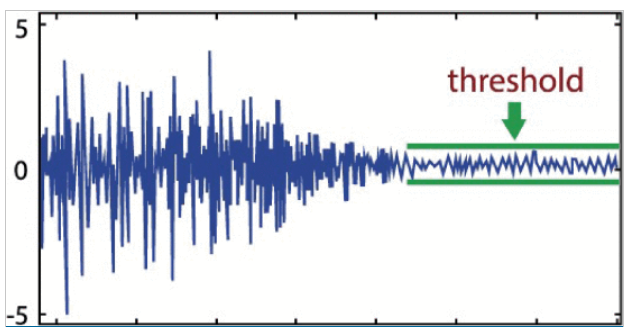
Other websites:

* <https://journals.physiology.org/doi/full/10.1152/japplphysiol.00636.2019>
* <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7242282/>
* <https://www.sciencedirect.com/topics/medicine-and-dentistry/facial-muscle>

Geneva minimalistic acoustic parameter set features:

* ~~Frequency related parameters~~ **UNRELATED TO EMG**
  + ~~Pitch~~
  + ~~Jitter~~
  + ~~Formant 1, 2 and 3 frequencies 🡨 Higher order frequencies of the fundamental frequency (the frequency at which your voice box vibrates)~~
  + ~~Formant 1 bandwidth~~
  + ~~Harmonics-to-noise ratio 🡨 Ratio of energy in the harmonics to energy outside the harmonics~~
* ~~Energy/amplitude parameters~~ **UNRELATED TO EMG**
  + ~~Shimmer~~
  + ~~Loudness~~
* Spectral parameters
  + **Alpha ratio** 🡨 Relevant if we can find a comparable threshold.
    - Ratio of summed energy from low-frequency period to high-frequency period
  + **Hammarberg index** 🡨 Relevant if we can find a comparable threshold.
    - Ratio of strongest energy peak from low-frequency period to high-frequency period
  + **Spectral slope** 🡨 Relevant
    - Linear regression slope
  + Formant 1, 2 and 3 relative energies
  + Harmonic difference H1-H2
  + Harmonic difference H1-A3
* Temporal features
  + Rate of loudness peaks.
  + Mean/std of voiced regions.
  + Mean/std of unvoiced regions.
  + Rate of voiced/unvoiced regions per second.

EMG Pattern Recognition in the Era of Big Data and Deep Learning:

* Page 3
  + Zero Crossings (ZC) - ?
    - <https://ieeexplore.ieee.org/document/8951119>
      * Number of times a signal crosses 0, usually has a threshold to avoid counting signal noise
      * 
      * There is no consensus for what value is the best threshold
    - <https://www.sciencedirect.com/science/article/abs/pii/S1746809418301447>
      * Study found that for surface EMG, zero crossings (ZC) and slope sign changes (SSC) required no threshold
      * Threshold of R = 0.1:1 needed for Willison amplitude (WAMP), myopulse percentage rate (MYOP) and cardinality (CARD)
  + Slope Sign Changes (SSC) - ?
  + Mean Absolute Value (MAV)
  + Waveform Length (WL)
  + Autoregressive Coefficients (AR)
  + Cepstral Coefficients (CC)
  + Willison Amplitude (WAMP)
  + Sample Entropy (SampEn)
* Page 5
  + Mean shift
  + Coordinates of center of gravity
  + Maximum values
* Page 6
  + “Recently, emotion recognition using multiple physiological modalities has gained attention as another important application that has benefited from the incorporation of surface EMG.”
* Page 7
  + Study by Healey and Picard – one of the earliest examples of multi-modal emotion recognition based on physiological signals
    - Used trapezius muscle and respiration
    - Measured stress levels
    - Used 24 subjects
    - Data from 17 subjects is publicly available
  + Other datasets
    - DEAP – Database for Emotion Analysis using Physiological signals
    - HR-EEG4EMO
    - DECAF
    - BioVid Emo DB
    - BioVid Heat Pain DB
    - Mentions some datasets have used the zygomaticus major and corrugator muscle
    - “To gain access to these datasets, the EULA must be printed, signed, scanned, and returned via email to the authors of each dataset. Upon approval, they will then provide a username and password that can be used to download the data.”
* Page 14
  + L-scale
  + Additional studies: [7, 8, 18-20, 22, 26, 28-30]
    - 7. Phinyomark, A.; Quaine, F.; Charbonnier, S.; Serviere, C.; Tarpin-Bernard, F.; Laurillau, Y. EMG Feature Evaluation for Improving Myoelectric Pattern Recognition Robustness. Expert Syst. Appl. 2013, 40, 4832–4840. [CrossRef]
    - 8. Tkach, D.; Huang, H.; Kuiken, T.A. Study of Stability of Time-Domain Features for Electromyographic Pattern Recognition. J. NeuroEng. Rehabilit. 2010, 7, 21. [CrossRef] [PubMed]
    - 18. Phinyomark, A.; Phukpattaranont, P.; Limsakul, C. Feature Reduction and Selection for EMG Signal Classification. Expert Syst. Appl. 2012, 39, 7420–7431. [CrossRef]
    - 19. Boostani, R.; Moradi, M.H. Evaluation of the Forearm EMG Signal Features for the Control of a Prosthetic Hand. Physiol. Meas. 2003, 24, 309. [CrossRef] [PubMed]
    - 20. Zardoshti-Kermani, M.; Wheeler, B.C.; Badie, K.; Hashemi, R.M. EMG Feature Evaluation for Movement Control of Upper Extremity Prostheses. IEEE Trans. Rehabilit. Eng. 1995, 3, 324–333. [CrossRef]
    - 22. Phinyomark, A.; Khushaba, R.N.; Ibáñez-Marcelo, E.; Patania, A.; Scheme, E.; Petri, G. Navigating Features: A Topologically Informed Chart of Electromyographic Features Space. J. R. Soc. Interface 2017, 14. [CrossRef] [PubMed]
    - 26. Phinyomark, A.; Khushaba, R.N.; Scheme, E. Feature Extraction and Selection for Myoelectric Control Based on Wearable EMG Sensors. Sensors 2018, 18, 1615. [CrossRef] [PubMed]
    - 28. Hudgins, B.; Parker, P.; Scott, R.N. A New Strategy for Multifunction Myoelectric Control. IEEE Trans. Biomed. Eng. 1993, 40, 82–94. [CrossRef] [PubMed]
    - 29. Khushaba, R.N.; Al-Timemy, A.H.; Al-Ani, A.; Al-Jumaily, A. A Framework of Temporal-Spatial Descriptors-Based Feature Extraction for Improved Myoelectric Pattern Recognition. IEEE Trans. Neural Syst. Rehabilit. Eng. 2017, 25, 1821–1831. [CrossRef] [PubMed]
    - 30. Phinyomark, A.; Scheme, E. A Feature Extraction Issue for Myoelectric Control Based on Wearable EMG Sensors. In Proceedings of the IEEE Sensors Applications Symposium (SAS), Seoul, Korea, 12–14 March 2018; pp. 1–6. [CrossRef]
  + Different EMG feature selection/projection methods - ?
    - Genetic Algorithm (GA)
    - Particle Swarm Optimization (PSO)
    - Ant Colony Optimization (ACO)
    - Principal Component Analysis (PCA)
    - Topological Data Analysis (TDA)
* Page 17
  + Spectrogram - ?

~~Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers:~~

* ~~Page 57~~
  + ~~Prosodic~~
    - ~~Pitch~~
    - ~~Energy~~
    - ~~Duration~~
  + ~~Spectral~~
    - ~~MFCC~~
    - ~~LPCC~~
    - ~~LFPC~~
    - ~~GFCC~~
    - ~~Formants~~
  + ~~Voice quality~~
    - ~~Jitter~~
    - ~~Shimmer~~
    - ~~HNR~~
    - ~~Normalized amplitude quotient~~
    - ~~Quasi open quotient~~
  + ~~Based on Teager energy operator~~
    - ~~TEO-FM-Var~~
    - ~~TEO-Auto-Env~~
    - ~~TEO-CB-Auto-Env~~
* ~~Page 58~~
  + ~~Discrete emotion theory 🡨 Assumes emotions can be divided into some combination of sadness, happiness, fear, anger, disgust, and surprise. We are analyzing the categories of sadness, happiness, fear, anger, disgust, and calm.~~
  + ~~Types of emotional recognition databases:~~
    - ~~Acted (simulated) emotion 🡨 Recorded by actors.~~
    - ~~Elicited (induced) emotion 🡨 Speakers have their emotions triggered.~~
    - ~~Natural speech emotion. 🡨 Audio from talk shows, recordings, etc.~~
  + ~~Preprocessing steps used in report:~~
    - ~~Framing 🡨 Speech signals are partitioned into fixed length segments.~~
    - ~~Windowing 🡨 Hamming window function~~
    - ~~Voice activity detection 🡨 Use endpoint detection to identify and extract both voiced and unvoiced speech.~~
      * ~~Zero crossing rate, short time energy and auto-correlation method are used to detect voiced/unvoiced speech.~~
    - ~~Normalization 🡨 z-normalization is used.~~
    - ~~Noise reduction 🡨 Minimum mean square error and log-spectral amplitude MMSE are the most successful.~~
    - ~~Feature selection and dimension reduction 🡨 Features are extracted, and a feature selection algorithm is used to remove redundant/least-useful features.~~
      * ~~A SVM classifier was able to get a 92.38 recognition rate with just 6 features.~~
      * ~~Luengo, I., Navas, E., Hernáez, I., Sánchez, J., 2005. Automatic emotion recognition using prosodic parameters. In: Ninth European Conference on Speech Communication and Technology.~~
* ~~Page 62~~
  + ~~Prosodic features 🡨 Intonation/rhythm, “para-linguistic” features (~~*~~We could calculate these, but this may not be useful as they specifically deal with syllables, words, phrases, and sentences in speech signals. We are using physiological signals, not speech signals~~*~~)~~
  + ~~Spectral features 🡨 Features extracted from the data when in a frequency domain.~~
    - **~~Mel Frequency Cepstral Coefficients (MFCC)~~** ~~🡨 Short-term power spectrum of speech signal~~
      * ~~To obtain:~~
        + ~~Utterances are divided into segments.~~
        + ~~Each segment is converted to the frequency domain using Fourier transform.~~
        + ~~Sub-band energies are calculated using a Mel filter bank.~~
        + ~~The logarithm of these sub-bands is taken.~~
        + ~~An inverse-Fourier transform is applied.~~
      * ~~Most widely used spectral feature~~
    - **~~Linear Prediction Cepstral Coefficients (LPCC)~~** ~~🡨 Shows differences with particular emotions.~~
      * ~~Can be obtained with a recursive method from Linear Prediction Coefficient (LPC)~~
    - **~~Log-Frequency Power Coefficients (LFPC)~~** ~~🡨 Measures spectral band energies using fast Fourier transform.~~
    - **~~Gammatone Frequency Cepstral Coefficients (GFCC)~~** ~~🡨 Similar technique to MFCC extraction, applies a Gammatone filter-bank instead of a Mel filter bank.~~
    - **~~Formants~~** ~~🡨 Amplitude peaks in the frequency spectrum, determines the phonetic quality of a vowel.~~
  + ~~Voice quality features 🡨 Determined by physical properties of the vocal tract.~~
    - **~~Jitter~~** ~~🡨 Variability of the fundamental frequency~~
    - **~~Shimmer~~** ~~🡨 Variability of the amplitude~~
    - **~~Harmonics-To-Noise Ratio~~** ~~🡨 Relative level of noise in the frequency spectrum of vowels~~
    - **~~Normalized Amplitude Quotient (NAQ)~~**
    - **~~Quasi Open Quotient (QOQ)~~**
    - **~~Difference between amplitudes of the first two harmonics (H1H2)~~**
    - **~~Maxima Dispersion Quotient (MDQ)~~**
    - **~~Spectral tilt/slope of wavelet responses (peak-slope)~~**
    - **~~Parabolic Spectral Parameter (PSP)~~**
    - **~~Shape parameter of the Liljencrants-Fant model of the glottal pulse dynamics (Rd)~~**
    - **~~Open Quotient Gradient~~**
    - **~~Glottal Opening Gradient~~**
    - **~~Skewness Gradient~~**
    - **~~Rate of Closure Gradient~~**
    - **~~Incompleteness of Closure~~**
* ~~Page 63~~
  + ~~Teager energy operator-based features 🡨 Features that depend on the Teager Energy Operator (TEO), used to detect stress in speech. Given a sampled speech signal x(n), the TEO is:~~
    - *~~We can calculate the TEO, but these features may not be relevant as we are using physiological signals rather than speech signals, so whatever we get as the TEO won’t represent the same thing~~*
    - ~~Features:~~
      * **~~TEO-Decomposed FM variation (TEO-FM-Var)~~**
      * **~~Normalized TEO Auto-Correlation Envelope Area (TEO-Auto-Env)~~**
      * **~~Critical Band Based TEO Auto-Correlation Envelope Area (TEO-CB-Auto-Env)~~**
    - ~~These features explore the variation in energy of airflow for voiced speech under stress~~
* ~~Page 64~~
  + ~~“While it is challenging to collect biosignals for both training and classification, they have the advantage of being uncontrolled […] people may hide their emotion in their speech; however, it is harder to alter biosignals.”~~
  + ~~Kim uses biosignals and speech to classify emotions:~~ **~~Kim, J., 2007. Bimodal emotion recognition using speech and physiological changes. Ro-bust speech recognition and understanding. InTech.~~**

Surface Electromyography Signal Processing and Classification Techniques

* Integrated EMG (IEMG)
* Mean Absolute Value (MAV)
* Modified Mean Absolute Value (MMAV1)
* Modified Mean Absolute Value 2 (MMAV2)
* Simple Square Integral (SSI)
* Variance of EMG (VAR)
* Root Mean Square (RMS)
* Waveform Length (WL)
* Willison Amplitude (WAMP)
* Log Detector (LOG)
* Slope Sign Change (SSC)
* Zero Crossing (ZC)
* Multi-Scale Amplitude Modulation-Frequency Modulation (AM-FM)

# References

The Geneva Minimalistic Acoustic Parameter Set

EMG Pattern Recognition in the Era of Big Data and Deep Learning

Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5800529/>

<http://ltu.diva-portal.org/smash/get/diva2:1557981/FULLTEXT01.pdf>

EMG Feature selection and classification using a Pbest-guide binary particle swarm optimization

* Mean absolute value (MAV)
  + Popular EMG feature widely used in EMG pattern recognition
* Wavelength (WL)
  + Frequently used EMG feature, represents cumulative length of waveform over time
* Root Mean Square (RMS)
  + Describes muscle force and non-fatigue contraction
* Maximum Fractal Length (MFL)
  + Recent EMG feature used to measure activation of low-level muscle contraction
* Average Power (AP)
  + Statistical feature for measuring energy distribution

OpenSMILE features (<https://audeering.github.io/opensmile/about.html#capabilities>):

* Audio-specific features (**MAY NOT BE APPLICABLE**):
  + Frame energy
  + Frame intensity / loudness
  + Critical band spectra (mel / bark / octave, triangular masking filters)
  + Mel / bark-frequency-cepstral coefficients (MFCC)
  + Auditory spectra
  + Loudness approximated from auditory spectra
  + Perceptual linear predictive (PLP) coefficients
  + Perceptual linear predictive cepstral coefficients (PLP-CC)
  + Linear predictive coefficients (LPC)
  + Line spectral pairs (LSP, aka LSF)
  + Fundamental frequency
  + Probability of voicing from ACF and SHS spectrum peak
  + Voice quality (jitter and shimmer) 🡨 Already discussed not using
  + Format frequencies and bandwidths
  + Zero and mean crossing rate 🡨 Was going to add this already
  + Spectral features
    - Arbitrary band energies 🡨 I think this means the sum of energies of certain ranges of frequencies, I think we already have this in the form of the twitch ratio
    - Roll-off points 🡨 Added
    - Centroid 🡨 **DONE (FOR SPECTRUM)**
    - Entropy 🡨 Added
    - Maxpos 🡨 **DONE**
    - Minpos 🡨 **DONE**
    - Variance (spread) 🡨 **DONE**
    - Skewness 🡨 Added
    - Kurtosis 🡨 **DONE**
    - Slope 🡨 **DONE**
  + Psychoacoustic sharpness spectral harmonicity
  + CHROMA (octave-warped semitone spectra) and CENS features (energy-normalized and smoothed CHROMA)
  + CHROMA-derived features for chord and key recognition
  + F0 harmonics ratios
* Video features (**NOT APPLICABLE**)
* Functionals
  + Extreme values and positions 🡨 **DONE**
  + Means 🡨 **DONE**
  + Percentiles and percentile ranges
  + Regression 🡨 **DONE (?)**
  + Centroid 🡨 **DONE**
  + Peaks
  + Segments
  + Sample values
  + Times/durations
  + Onsets/offsets
  + Discrete cosine transformation (DCT)
  + Zero crossings
  + Linear predictive coding (LPC) coefficients and gain