To identify a correlation between IMU and microphone data in earable computing with regards to chewing

Gregor Figueira Comojo

Supervisors: Przemysław Pawełczak, Vivian Dsouza

EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 25, 2023

Name of the student: Gregor Figueira Comojo
Final project course: CSE3000 Research Project
Thesis committee: Przemysław Pawełczak, Vivian Dsouza, Hayley Hung

Abstract

This research explores the correlation between chewing activities and non-chewing activities using an Arduino microcontroller. Chewing samples are recorded by attaching the microcontroller to the back of the jaw, underneath the ear. The microcontroller collects data from a microphone capturing audio data of chewing sounds, as well as an Inertial Measurement Unit (IMU) collecting motion and orientation data of jaw movements. The collected data is processed using signal processing techniques, extracting relevant features from the microphone data, such as intensity, frequency content, and features related to acceleration, orientation, and jaw movement patterns from the IMU data. Statistical analysis, employing correlation metrics like Pearson correlation coefficient and Spearman's rank correlation coefficient, determines the correlation between the extracted features from the microphone and IMU data. Conclusions from the analysis indicate that pre-processing and feature extraction techniques are needed to establish meaningful correlations between the IMU and microphone data. The sliding window approach shows promising results, particularly in correlating the sum of energy from the audio with the sum of gyro data, specifically in the y- and z-axes. The accelerometer data does not exhibit significant correlations, but it can be useful as a threshold for detecting the start of chewing events based on zero crossings. Furthermore, the findings reveal that food texture and density play a larger role than anticipated in determining the correlation between chewing patterns and sensory data. The research outcomes contribute to various fields, including dentistry, nutrition, and human-computer interaction.

1 Introduction

Mobile computing has made a big impact in humanity, especially since the introduction of wearable computing. A wearable computer is a computing device worn on the body, for instance a watch, a chain, glasses or a earphone. [1] Turning these devices into smart devices can give us valuable information, allowing for better insight into our life. The smartwatch is a great example for this, as it can monitor everything from the user's activity to their heart rate. [2]

Research for earable computing focuses on the possibilities for gathering information from an earphone. These allow capturing new kinds of information on a regular basis, due to the proximity of the sensors to the eyes, nose, mouth and ears. This information can be very helpful in fields like healthcare, as it might be able to detect eating patterns.

This paper will focus on identifying the correlation between the inertial measurement unit (IMU) and the microphone data with regards to chewing. Other papers suggest that there is a correlation between the two sensors as using more than one sensor yields better results. [3] However, the correlation itself is never researched. Identifying the correlation may increase the accuracy of detecting chewing due to the fact that better features from the data can be extracted.

Identifying the correlation leads to a better understanding of the sensor data which progresses the accuracy of detecting chewing. Increasing the accuracy of detecting one activity means that the inaccuracy of detecting other activities decreases, speech detection for example. As mentioned earlier, the found correlation allows us to define better features. These can, for example, be used to train a tiny machine learning model which can be stored in the earphone. Due to the limitation of computational power and storage, it is always beneficial to reduce the use of resources.

The upcoming sections will go through prior research on earable computing, the methodology used to answer the research question and the results of the research. The paper is closed with a conclusion on the results and a discussion with regards to possible improvements and optional further research questions from these findings.

2 Method

This research aims to explore the correlation between chewing patterns and sensory data by utilizing an Arduino microcontroller. The microcontroller will be attached to the back of the jaw underneath the ear to record chewing samples. A microphone will capture audio data of the chewing sounds, while an IMU will collect motion and orientation data of the jaw movements.

The Arduino platform is used to collect data from the IMU and microphone simultaneously. The recorded data will be stored in the microcontroller's memory and a serial connection will be established on the COM port for convenient retrieval. To analyze the collected data, relevant features will be extracted from the microphone data, including measures of intensity, frequency content, and spectral characteristics. Similarly, features related to acceleration, orientation, and jaw movement patterns will be extracted from the IMU data. In order to process the sensor data effectively, signal processing techniques will be utilized which are explained in-depth in Section 3.5.

Statistical analysis will be performed to calculate the correlation between the extracted features from the microphone and IMU data. Established correlation metrics such as the Pearson correlation coefficient or Spearman's rank correlation coefficient will be used. [4] The obtained correlation coefficients will be compared across different features, for which some are from [5], to determine if certain features demonstrate a stronger relationship.

The strength and direction of the correlations will be analyzed to gain insights into the relationship between chewing patterns and sensory data. Additional exploratory analysis will be conducted if necessary to identify potential factors influencing the observed correlations. The research findings will contribute to various fields, including dentistry, nutrition, and human-computer interaction.

3 Experimental Setup

The experiment is crafted in such a way that it is easily reproducible. In this particular case the Arduino Nano BLE 33 Sense is used and the source code is configured to this microcontroller. However, the experiment could be reproduced with any other microcontroller or embedded device with similar specifications. This refers to the IMU and microphone having similar specifications with regards to the sample rate and sensitivity.

3.1 Hardware Specifications

The experiment was performed on an Arduino Nano 33 BLE Sense. The microcontroller was built on top of the nFR52840 processor, which has a clock speeds of 64MHz and access to 256 KB SRAM. The use of pins was not needed as the required sensors were already built in. The microphone is a MP34DT05, it is an omnidirectional microphone. Ideal for this experiment as the sound source, the mouth, is not directly aimed at the microphone. With the 16kHz sample rate and 16-bit samples, the audio it records is high definition on playback and distinguishing between chews and other noises was rather easy by human ear. For motion detection, the Arduino has an IMU (LSM9DS1) built in allowing access to three forms of data. Accelerometer, Gyroscope and magnetometer. In this experiment this experiment only the former two were used. The accelerometer is an electromechanical sensor used to measure acceleration forces in three-dimensional space. These forces are expressed as gravitational constants (g), ranging between [-4, +4]g. The gyroscope measures the angular velocity in three-dimensional space. The velocity is measured in degrees per second (dps) and the domain of the sensor is [-2000, +2000] dps. The sample rate for all the sensors on the IMU is 104Hz[1].

3.2 Environment Setup

Before the data can be recorded, the microcontroller needs to be attached. The OpenEarable paper [6] works with an earable that goes around the ear such that the IMU has the possibility to reach behind the jaw as seen in Figure 1. In this case, the Arduino was put alongside the jaw such that the IMU recorded a lot of motion. The axis of the of the IMU are displayed in the figure, the positive y-axis goes out of the paper. It is worth noting that IMU results might be different with other microcontrollers due to the shape of the controller and/or the orientation of the IMU. For sticking the Arduino to the back of the jaw, Leukopor tape was used as it has the right amount of adhesive and causes minimal to no pain upon removal. However, any other tape or adhesive could be used. A serial connection to a laptop was used to collect the sensor data, so for practicability reasons the Arduino was oriented such that the USB-port aimed downwards, making connecting it more feasible.

[1]More specifications at https://docs.arduino.cc/hardware/nano-33-ble

Figure 1: Arduino attachment location on the jaw

3.3 Data Collection

This section is split into two different parts. Firstly the technical aspects of recording the samples on the Arduino is explained. After that, the actual recording of the samples is discussed such that these could be accurately reproduced.

Technical Aspects

As mentioned earlier, the data transfer happens through a serial connection. With a fast calculation it shows that the 256KB SRAM on the microcontroller would be able to hold $256000/(16000 \cdot 2) = 8$ seconds of microphone data in it's memory without taking the used memory in account by the OS and the program itself. This is not very useful for recording samples as chewing happens often in longer periods of time. The IMU, on the other hand, generates much less data in comparison. The IMU data is stored as UTF-8 characters with 2 decimal accuracy. Meaning that the accelerometer, in the worst case scenario[2], generates $5 \cdot 3 \cdot 104 = 1560$ bytes per second and the gyroscope generates[3] $8 \cdot 3 \cdot 104 = 2496$ bytes per second. So both sensors together generate $1560 + 2496 = 4056$ bytes per second and thus $256000/4056 = 63.12$ seconds of IMU data can be stored. This is plenty as the recorded samples are between 10s and 20s.

Due to the high frequency of microphone data and thus the short time span the RAM can hold, the microphone data is sent out in real time. The IMU data is stored in the RAM and every 10s one packet containing all the data is sent over serial. The process of collecting a 10s sample is to record 20s of audio, this ensures that the there will be at least one full cycle of 10s where the audio data aligns with the 10s packet

[2]The number -1.23 uses 5 UTF-8 characters

[3]The number -1234.56 uses 8 UTF-8 characters

of IMU data. For a 20s sample the recording time would be 40s.

Recording Process

The Arduino was attached to the jaw as seen in Figure 1. Four distinct activities were recorded for one minute. The activities included talking, texting, eating peanuts and eating a cucumber. One sample contains both talking and chewing, this is for plotting purposes. The talking sample contains a passage from a book read out loud. Texting is one minute of the subject using their phone, so it contains multiple on-phone activities such as swiping or scrolling. For all activities the voice is not used. In both eating samples the subject aimed not to move their head unnecessarily such that mostly jaw movement was recorded. The plot of such a sample is visible in Figure 2. This sample is not used for calculating any correlation but for visualizing the difference between chewing and talking. It starts with talking and transitions to chewing at around sample 70,000 (sample 500 for IMU data). Then, at sample 380,000 (sample 2,600 for IMU data), it reverts back to talking. The transitions are marked by the vertical red lines.

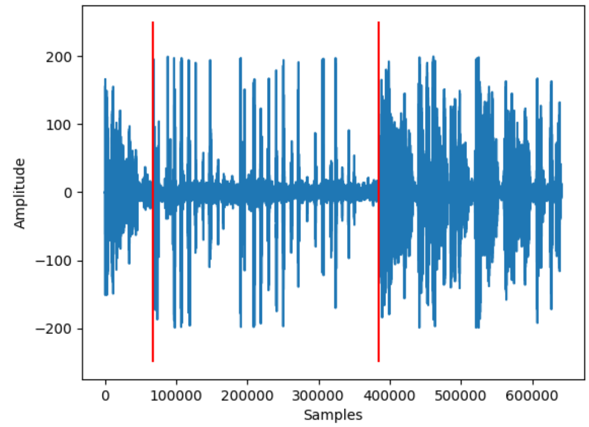
3.4 Data Acquisition

The computer receives the serial packets and stores them into a file using a Python script. The *PySerial* library allows for simple code to accomplish this. Another python script separates the microphone data from the IMU data and stores these into different files. The fact that the microphone is stored as two's compliment integers and the IMU as character allows the use of pattern recognition. The IMU data consist of lines of three floats from the accelerometer and three floats from the gyroscope separated by a tab character between the two. So a mathod for identifying IMU data is looking for lines that are shorter or equal to 39 bytes and contain a tab character. The presence of three "." characters on both sides of the tab indicates the presence of IMU data. This approach may have some false positives but the likelihood is minimal. If a false positive occurs, the maximum amount of audio data lost is 19 samples.

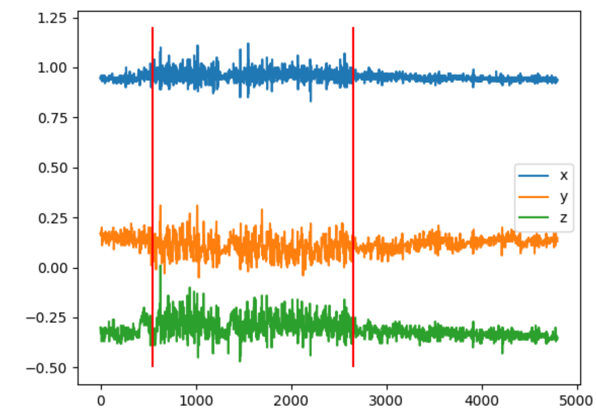
3.5 Data Processing

The data is pre-processed to remove noise and corrupt data. Sensor data often contains faulty measurements. The following paragraph will glance over the pre-processing strategy, afterwards different types of processing that were applied will be described.

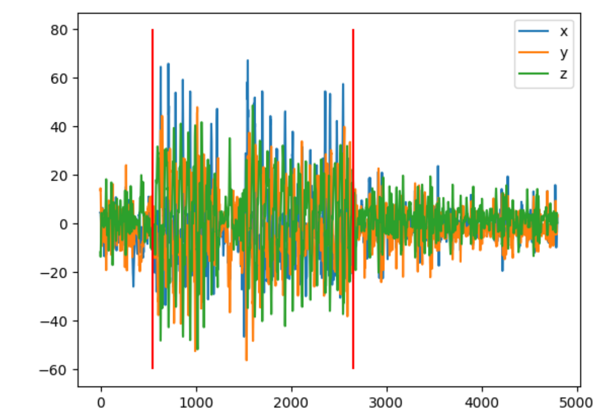
Hardware sensors are prone to corrupt data points. Points like these will most likely influence the results when ignored, thus it is of utmost importance to cleanse sensor data. Cleansing is the process of detecting and correcting the faulty data points. In the case of the microphone data, outliers, so called spikes, were recorded as can be seen in Figure 3a. By applying a filter that gets rid of the extreme outliers which is a lot more balanced as seen in 3b. The filter interpolates between the preceding 5 samples below the threshold and the subsequent 5 samples for each outlier. With this filter the audio contains no clicks or pops on playback. However, it is very unlikely that unprocessed audio, think of audio processing like compression or limiting, produces a waveform like



(a) PCM audio data of interleaved talking and chewing activities



(b) Accelerator data of interleaved talking and chewing activities



(c) Gyroscope data of interleaved talking and chewing activities

Figure 2: Microphone data plots

seen in Figure 3b. [7] Adjusting the threshold to 200 results into Figure 3c, this is what is expected for microphone data to look like. Although Figure 3b and Figure 3c look vastly different, there is no detectable difference in the audio playback.

With the cleansed data the statistical analysis can be performed. The methods used for the analysis are the Pearson coefficient and Spearman's rank and Kendall's Tau.

Firstly, the correlation between the sum of the energy and the IMU was calculated. The sample rates of both sensors do not correspond, with the microphone data containing approximately $16000/104 = 153.84$ samples per single IMU sample. Summing up all these audio samples, yields a data set that corresponds in frequency. This opens up the possibility to mathematically calculate the correlation between the sum of energy and the IMU data.

Secondly the correlation between the sum of energy per frequency band and the IMU was calculated. Just as in the first calculated, the sample rates were synchronized by using the Fourier Transform. This transform allowed for adding up the amplitudes in their respective band and getting more in-depth correlation matrices. This method is inspired by the method from the paper [5] were they used a similar approach to detect chewing only from microphone data.

The sum of the gyroscope was also extracted and used for identifying a correlation. For this feature was the sliding window approach additionally used to see how different time frame can influence the results.

The source code for the Arduino and all scripts can be found in the project repository [LINK]

4 Research Results

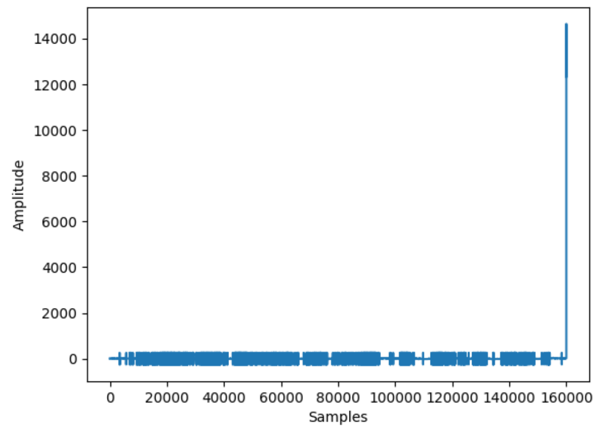
Before the results are presented, the explanations of the Pearson's correlation coefficient and the Spearman's rank correlation coefficient are given. These are the two metrics used to calculate the correlations throughout this whole segment.

The Pearson's correlation coefficient is a statistical measure that quantifies the linear relationship between two continuous variables. It assesses the strength and direction of the linear association, ranging from -1 (perfect negative correlation) to +1 (perfect positive correlation), with 0 indicating no linear correlation.

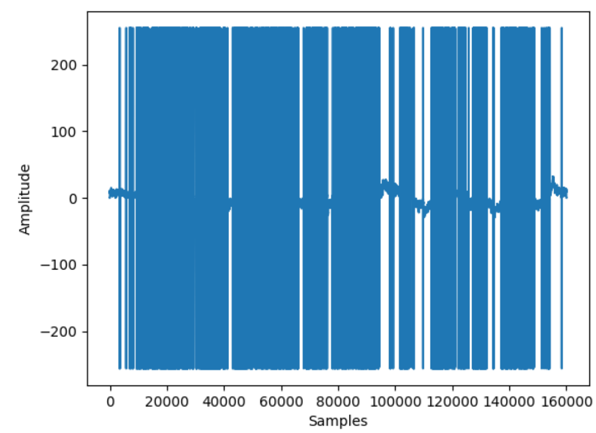
Spearman's rank correlation coefficient, on the other hand, is a non-parametric measure of the monotonic relationship[4] between two variables. It assesses the similarity in the relative order or ranking of the variables rather than the specific linear relationship. It ranges from -1 to +1, where -1 indicates a perfect decreasing monotonic relationship, +1 indicates a perfect increasing monotonic relationship, and 0 indicates no monotonic relationship. [8]

This section contains methods that yielded good and bad results. Table 1 shows the Pearson's correlation coefficient for every activity between the sum of audio energy and the raw IMU data. No matter what activity there is no correlation between the sum of energy and any of the IMU data. Thus this approach is discarded. Table 1 displays the Spearman's coefficients, this was calculated in the case that the sensors

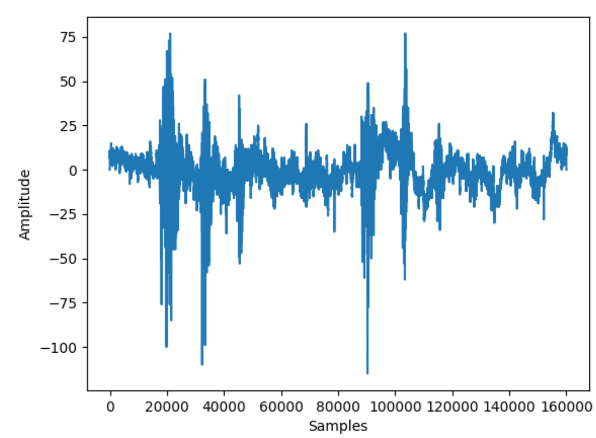
[4]A monotonic relationship is a relationship between ordered data



(a) Microphone data without filtering



(b) Microphone data with values were the absolute value is smaller than 500 amplitude



(c) Microphone data with values were the absolute value is smaller than 500 amplitude

Figure 3: Microphone data plots

were not correlated in a linear manner. However by looking at the result it can be concluded that there is no correlation between these features.

Table 1: Pearson's Correlation Between Audio Energy and IMU Data For All Activities

	Peanuts	Cucumber	Talking	Texting
Acc_x	0.006	0.010	0.001	0.071
Acc_y	-0.019	-0.024	-0.022	-0.130
Acc_z	0.020	-0.031	-0.006	-0.013
$Gyro_x$	-0.006	0.035	0.014	-0.015
$Gyro_y$	0.005	0.030	-0.005	0.045
$Gyro_z$	0.045	-0.022	0.017	-0.005

Table 2: Spearman's Correlation Between Audio Energy and IMU Data For All Activities

	Peanuts	Cucumber	Talking	Texting
Acc_x	-0.024	-0.062	-0.032	-0.011
Acc_y	-0.049	-0.016	-0.014	-0.004
Acc_z	-0.075	-0.133	-0.077	-0.024
$Gyro_x$	-0.081	-0.156	0.058	0.018
$Gyro_y$	0.012	-0.062	0.011	0.017
$Gyro_z$	0.022	0.046	-0.024	0.010

Afterwards, the correlation between the sum of energy and the sum of IMU data, meaning the sum of the accelerator and the sum of the gyroscope, was calculated and is displayed in Table 3. This in itself does not yield any better results whatsoever, so a sliding window approach was used. All values inside the window were summed up, this essentially meant that the time frame for which the correlation will be calculated is bigger. Correlations were calculated from window sizes between 10 and 1000 as to account for the uncertainty regarding the optimal window size. Figure 4 displays the correlations for eating a cucumber and peanuts for the different chosen window sizes. Since the sample frequency of the IMU is 104Hz, calculating the time span of the window size $windowsize/104$ seconds, which approximates 1 second per 100 window size.

Table 3: Spearman's Correlation Between Audio Energy and Sum of IMU Data For All Activities

	Peanuts	Cucumber	Talking	Texting
Acc_{sum}	-0.027	0.012	0.073	0.007
$Gyro_{sum}$	-0.038	-0.075	0.097	-0.017

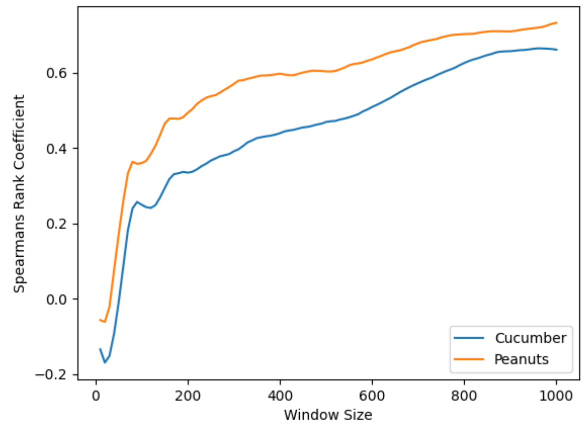


Figure 4: Spearman's Rank Between Sum of Audio Energy and Sum of Gyroscope for Different Window Sizes

The correlation gets stronger when the window size is increased. This also corresponds with the low values in Table 2 as these can be considered as having a window size of one. This is a positive finding however, since identifying the correlation has to goal to make detecting chewing activities easier, it is also important to compare the found correlation with the correlations for talking and texting. Figures 6a and 6b show the difference in correlations from the activities, this is calculated by simply subtracting the non-chewing activity. It suggests a window size of approximately 300 for finding a solid difference in correlation when compared to talking. Figure 5 shows that for all windows sizes does not yield any significant difference in correlations and can thus be neglected.

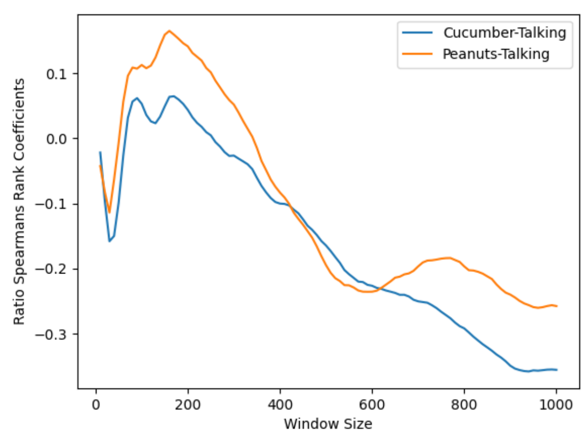


Figure 5: Difference between Spearman's Rank of Chewing and Talking for Different Window Sizes For Gyroscope X data

However, this can differ for each person. Figure 6c displays the difference between correlations for another participant. The orange line represents the different between chewing and talking from participant 2 (P2), the blue line represent the difference between chewing from P2 and talking from P1.[5] The optimal window to maximize the difference between correlations is approximately 500, which is significantly higher than the window for P1. Besides that, it is very interesting to see that by comparing correlations from chewing from P2 and talking by P1 there are no significant differences. The environment in which the samples were recorded was different, the microphone samples from P2 contained more background noise than those of P1. So it can be concluded that having background noise does not yield any problems when it stays consistent throughout the samples.

The conclusion for texting is different, both correlations are about the same but this does not yield any problems for detecting chewing. The gyroscope barely measures any signal when texting as seen in Figure 7, especially compared to chewing (see Figure 2c) and thus a threshold could eliminate falsely detecting texting as chewing.

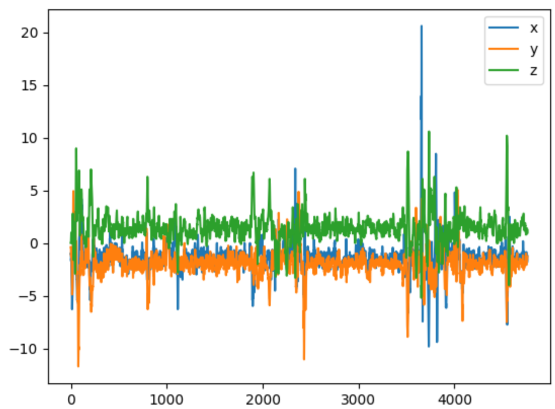
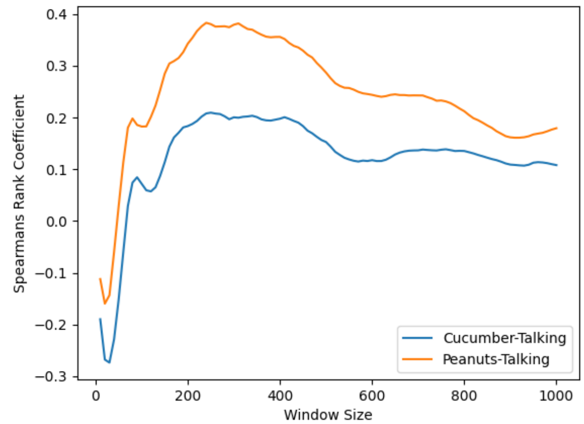


Figure 7: Gyroscope Data of Texting

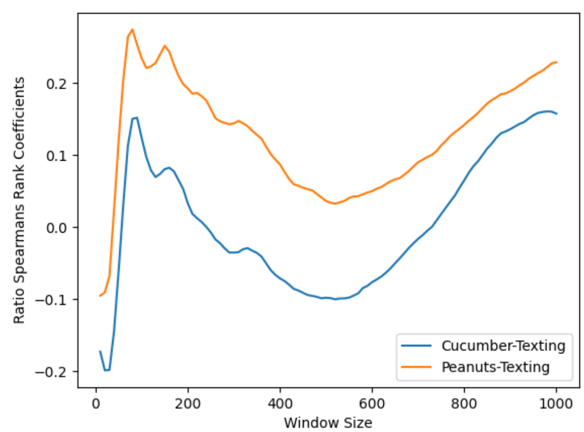
Another finding in the correlation is the difference between foods. The results indicate that the cucumber correlation was for all window sizes weaker than the correlation of eating peanuts. An explanation for this is the fact that peanuts are more dense than cucumbers, but they are less noisy. The IMU signal for is a little bit stronger for peanuts than for chewing cucumbers however, the sound energy for cucumbers is marginally larger. This is most likely the cause of the difference in the correlation coefficients.

Figures 8a and 8b show the correlation between the sum of gyroscope and the sum of different frequency bands of the audio energy. To get these values per frequency band, the Fourier Transform was used. These graphs display that the correlation is not stronger when the lower or higher frequencies are isolated compared to the full frequency spectrum. Avoiding the Fourier transform is advantageous for embedded systems due to its high computational cost, making this finding a positive one.

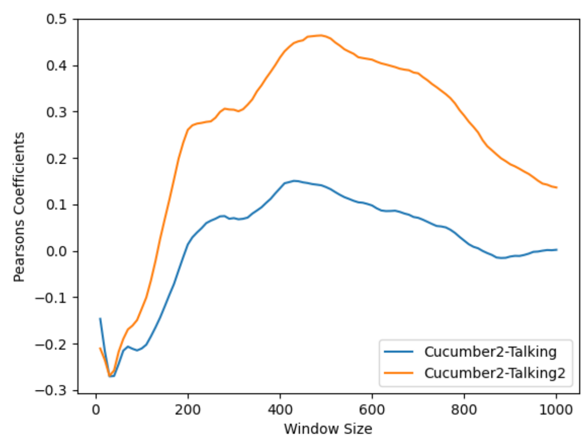
[5]There is only cucumber data as the chewing data for eating peanuts was corrupt



(a) Difference between Spearman's Rank of Chewing and Talking for Different Window Sizes

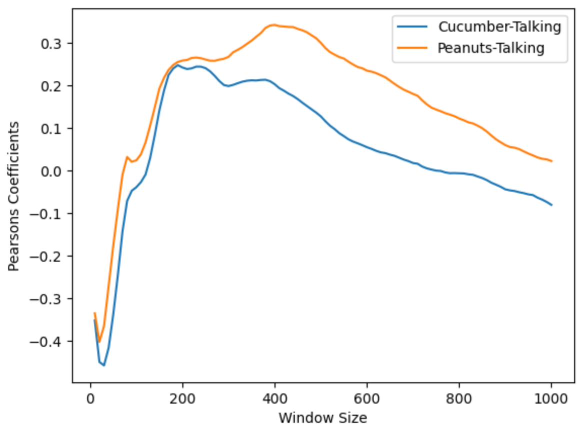


(b) Difference between Spearman's Rank of Chewing and Texting for Different Window Sizes

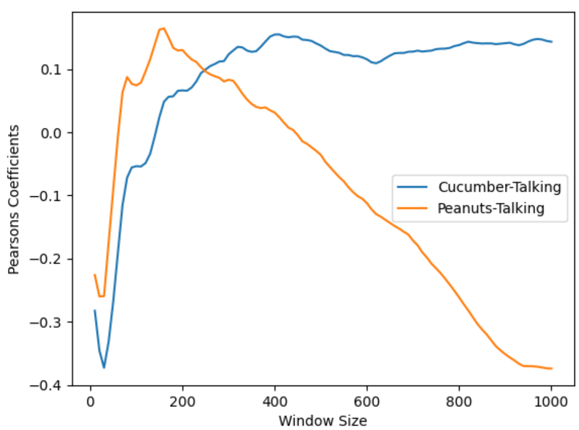


(c) Difference between Spearman's Rank of Chewing and Talking from different participants

Figure 6: Difference between Spearman's Rank correlations over different data samples



(a) Fourier Transform under 1000Hz



(b) Difference between Spearman's Rank of Chewing and Texting for Different Window Sizes

Figure 8: Fourier Transform above 1000Hz

Zero crossings have been calculated and are demonstrated in the Tables 4 and 5.[6] It is noticeable that the zero crossings for the accelerometer are higher for chewing activities than talking and texting for both participants. This can be crucial for detecting chewing by a threshold. There is no clear correlation between accelerometer and gyroscope crossings with respect to chewing, although the missing data of P2 eating peanuts might have given more insight.

Table 4: IMU Zero Crossings P1

	Peanuts	Cucumber	Talking	Texting
Acc_x	545	617	169	9
Acc_y	264	323	171	8
Acc_z	447	518	248	17
$Gyro_x$	333	455	697	710
$Gyro_y$	343	396	636	717
$Gyro_z$	252	264	479	493

Table 5: IMU Zero Crossings P2

	Peanuts	Cucumber	Talking	Texting
Acc_x	x	485	120	71
Acc_y	x	366	48	48
Acc_z	x	548	193	43
$Gyro_x$	x	567	517	540
$Gyro_y$	x	602	603	617
$Gyro_z$	x	518	477	447

Root Mean Square (RMS) values for the samples are displayed in the Tables 6 and 7. First of all, the RMS for accelerometer for all the samples are low and will not be taken into consideration. The RMS for the gyroscope for chewing activities is for P1 almost three times higher than talking but, the RMS for the microphone is approximately halve for chewing activities. This ties in with the conclusions found for the microphone energy and the gyroscope sums. The texting activities from P1 and P2 show different magnitudes of values as P2 moved their head a lot more than P1.

Table 6: RMS values for P1

	Peanuts	Cucumber	Talking	Texting
Acc_x	0.02	0.02	0.01	0.01
Acc_y	0.07	0.04	0.03	0.01
Acc_z	0.06	0.04	0.03	0.01
$Gyro_x$	14.35	13.44	4.33	1.40
$Gyro_y$	12.59	12.77	5.20	1.35
$Gyro_z$	14.59	14.73	5.53	1.42
Mic	149.73	172.71	337.37	211.09

[6] To ensure regularity, the column for peanuts was included

Table 7: RMS values for P2

	Peanuts	Cucumber	Talking	Texting
Acc_x	x	0.03	0.02	0.03
Acc_y	x	0.07	0.05	0.08
Acc_z	x	0.06	0.02	0.09
$Gyro_x$	x	14.84	5.77	14.47
$Gyro_y$	x	12.00	6.63	9.73
$Gyro_z$	x	7.26	5.89	7.41
Mic	x	146.76	401.35	160.04

5 Responsible Research

The earable device used in the study involves wearing a microphone and IMU that records the user's activities and environment. However, the recorded samples do not contain any private information or indicators to private information. The samples are made public for reproducibility purposes and the participants were made aware of that at the time of recording.

Another ethical concern pertains to the sensitive nature of the data collected by the device. Even though the current research only focuses on correlations, the bigger pictures is detecting chewing for health reasons. Calculating health metrics such as chewing patterns, food intake, and food types exposes personal information. This data could be exploited by malicious actors for targeted advertising or by medical insurance companies for patient classification.

It is important to note that the results and conclusions presented in the research are based on analyzing audio data from only two individuals and limited food samples. Results may vary among individuals with different eating habits and recording environments. The recordings used for calculating the correlations are available in the TUDelft GitLab repository, allowing others to reproduce the results.

6 Conclusions and Future Work

In conclusion, the analysis of the data reveals several insights. Firstly, the unprocessed IMU data alone does not provide meaningful results in terms of correlation. This suggests that additional pre-processing or feature extraction techniques are necessary to any sort of correlation between the IMU and microphone data. Calculating the sum of energy or the RMS produces good results.

Secondly, the sliding window approach yields solid results with regards to a correlation between the sum of energy from the audio and the sum of the gyro data. When calculating the sum of the gyro data, the x-axis did not provide any significant improvement and can be left out of the equation. The y- and z-axis are the correlating factors. For differentiating between talking and chewing a sliding window size 350 samples seems to be the sweet spot for one participant. The peak in the difference between correlations can differ for other participants and is not necessarily a given. Calculating the correlation for different frequency bins by using the Fourier transform did not change this finding.

Thirdly, the accelerometer did not correlate in any significant manner for all the features that were extracted. However, it can be useful as a threshold when calculating the zero crossing. Chewing events produce a lot more zero crossings than talking or texting and might be a good indicator for detecting the start of a chewing event.

Lastly, the difference in correlations between foods. The results demonstrate that the texture of different foods influence the correlation a lot more than initially anticipated. The hypothesis was that quieter chews would be harder to detect however, the relationship between robustness and the loudness of the food influences the correlation. Food texture and density play a larger role in finding a correlation than anticipated.

7 Discussion

In future research, the correlation between more food could be examined as here only two types of food were recorded. The control data was also not complete, the peanuts sample of P2 got corrupt and made it harder to conclude without making assumptions.

The sampling and/or the cleansing of the data might also need some reworking. In general the interpolating over the outliers yielded solid results however, in some cases there were still spikes in the data. Even though this might be minuscule, they still influence the results in an uncertain way.

The attachment of the Arduino could also have a influence on the data. In future research the use of the Bluetooth functionality might be beneficial as the cable weights down the Arduino. This could distort the IMU signal as the jaw movements are proportionally small and thus a relatively heavy cable could skew the data.

The cause of the shift in peaks for the Spearman's rank difference between chewing and talking, could be researched more. As of now it is not clear if this has to do with the chewing speed or with the environment change.

References

[1] Mads Soegaard and Rikke Friis Dam. *Encyclopedia of Human-Computer Interaction*. Interaction Design Foundation, 2013.

[2] Dung Phan, Lee Yee Siong, Pubudu N. Pathirana, and Aruna Seneviratne. Smartwatch: Performance evaluation for long-term heart rate monitoring. In *2015 International Symposium on Bioelectronics and Bioinformatics (ISBB)*, pages 144–147, 2015.

[3] Roya Lotfi, George Tzanetakis, Rasit Eskicioglu, and Pourang Irani. A comparison between audio and imu data to detect chewing events based on an earable device. In *Proceedings of the 11th Augmented Human International Conference*, AH '20, New York, NY, USA, 2020. Association for Computing Machinery.

[4] Martina Udovičić, Ksenija Baždarić, Lidija Bilić-Zulle, Mladen Petrovečki, et al. What we need to know when calculating the coefficient of correlation? *Biochemia Medica*, 17(1):10–15, 2007.

[5] Sebastian Päßler and Wolf-Joachim Fischer. Food intake monitoring: Automated chew event detection in chewing sounds. *IEEE Journal of Biomedical and Health Informatics*, 18(1):278–289, 2014.

[6] Tobias Röddiger, Tobias King, Dylan Ray Roodt, Christopher Clarke, and Michael Beigl. Openearable: Open hardware earable sensing platform. In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers*, UbiComp/ISWC '22 Adjunct, page 246–251, New York, NY, USA, 2023. Association for Computing Machinery.

[7] Earl Vickers. The loudness war: Background, speculation, and recommendations. *Journal of The Audio Engineering Society*, 2010.

[8] OpenAI. Chatgpt. https://openai.com, 2021. Prompt: "Explain in short what the Pearson's and Spearman's rank coefficient are".