Earable Computing: What can we learn from sensors close to the ear?

Cristian Cutitei

ccutitei@student.tudelft.nl

Proffessor: Przemysław Pawełczak

Supervisor: Vivian Dsouza

2005 2015 2018 2025

Introduction

The World Health Organization (WHO) has reported a significant increase in nutrition-related diseases like obesity and diabetes in recent centuries. This is primarily due to factors like a sedentary lifestyle and irregular eating habits with high calorie intake. While health wearables have gained popularity for addressing these issues, they mainly focus on monitoring physical activity and fail to address the second half of the problem, which is precise nutritional intake measurement.

So what is there to do? Simple, just monitor the eating habits as well.

Objective

That souned easy enough, but how do we do that? We can not just stick sensors in the human mouth, throat and stomac and gather data every time food intake is detected. Way too invasive, and contrary to other wearable devices, not very compact and comfortable, or, dare we say stylish.

Food intake is done through the food tract, how can it be monitored elsewhere ? Yes, it is, but there are more ways of sensing intake, one useful characteristic is sound. What for others is an annoying and potentially disgusting trait of food consumption, is the key to this problem. By algorithmically looking at the sounds produced, automatic detection is not only possible but not so resource intensive, computationally speaking. This means an embedded, low-power system, located such that it can easily pick up sounds from the mouth is enough to solve this monitoring problem.

Related literature

Of course, this is not the first time somebody taught to do this exact very thing. As far back as 1963, Drake investigated the acoustics of chewing hard, crisp food. Several research groups developed devices to record food intake sounds from his findings.

Later, Amft investigated different possible application places at the human head for their usability for food intake sound recording and found the ear canal to be the most suitable place.

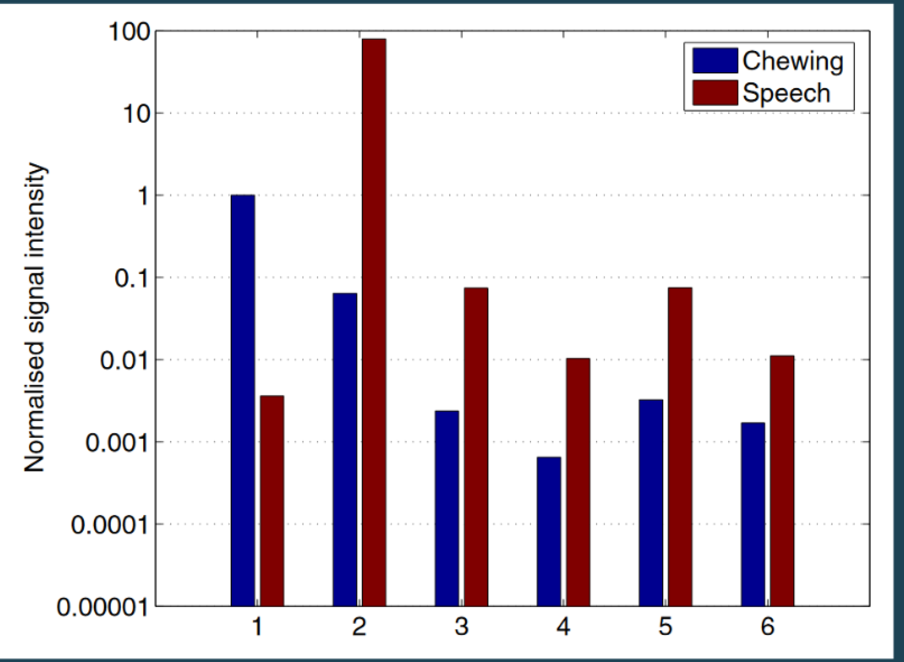
Analysis

These is the data that drove the findings and the conclusion of this research project

	Maximum Sound Energy	Low-Pass Filter	Maximum Slope
Apple-1	0.73	0.78	0.74
Apple-2	0.69	0.74	0.72
Carrot-1	0.72	0.77	0.70
Carrot-2	0.73	0.77	0.70
Chip-1	0.70	0.78	0.72
Chip-2	0.70	0.79	0.72

This tables values represent the procentages of accuracy of food detection based on the food consumed and algorithm.

Microphone	Position
1	Inner ear, directed towards eardrum (Hearing aid position)
2	2cm in front of mouth (Headset microphone position)
3	At cheek (Headset position)
4	5cm in front of ear canal opening (Reference position for audible chewing sounds)
5	Collar (Collar microphone position)
6	Behind outer ear (Hidden by the outer ear, used by older hearing aid models)



These depict the ratio of the amplitudes of chewing sounds and speech based on the position of the microphone

Methodology

Reading the Related Literature is the first step in every research project as it provides insight into what peers have already found. No need to reinvent the wheel. These are the steps I took:

- Separated findings into what is useful
- Organised algorithms already created and implemented them locally for testing
- Created a dataset specifically for testing the accuracy of each algorithm
- Gathered results
- Improved on their accuracy and effectiveness
- Tested system in new nevironments and assesed results

Results/Findings

Contrary to the current craze regarding machine learning and AI, they are not always the solution, especially on an embedded system. According to the results of this research and others, the accuracy from using a machine learning approach offers few improvements to the accuracy and sometimes even falls behind other seemingly simple algorithms. This, coupled with how computationally heavy they can become, they were out of the question.

The microphone's location used for recording was also in question when conducting this research. The best location was inside the ear canal. From amplitude analysis, it was easy to see that from this location, the chewing sound intensity was significantly larger than that from speaking.

Conclusion

The rising numbers of diseases related to food intake (e.g. obesity) and eating disorders have become a growing concern. To counteract this trend, a proposed solution is to have an automated food intake analysis. Three algorithms were created and tested against real-world data to achieve this task. An additional algorithm was created to reduce the chewing count detected in speech. Still, the algorithms fail to separate chewing audio from the rest and generate too many insertions where chewing is absent.