

Tracking and Counting Motion for Monitoring Food Intake Based-On Depth Sensor and UDOO Board: A Comprehensive Review

Muhammad Fuad bin Kassim¹, Mohd Norzali Haji Mohd^{1,2}

¹Embedded Computing Systems (EmbCos), Department of Computer Engineering, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia

²Modeling and Simulation (BIOMEMS) Research Group, Department of Electronic Engineering, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia

Corresponding author: fuadkassim93@gmail.com, norzali@uthm.edu.my

Abstract. Technology is all about helping people, which created a new opportunity to take serious action in managing their health care. Moreover, Obesity continues to be a serious public health concern in the Malaysia and continuing to rise. Obesity has been a serious health concern among people. Nearly half of Malaysian people overweight. Most of dietary approach is not tracking and detecting the right calorie intake for weight loss, but currently used tools such as food diaries require users to manually record and track the food calories, making them difficult for daily use. We will be developing a new tool that counts the food intake bite by monitoring hand gesture and face jaw motion movement of caloric intake. The Bite count method showed a good significant that can lead to a successful weight loss by simply monitoring the bite taken during eating. The device used was Kinect Xbox One which used a depth camera to detect the motion on person hand and face during food intake. Previous studies showed that most of the method used to count bite device is worn type. The recent trend is now going towards non-wearable devices due to the difficulty when wearing devices and it has high false alarm ratio. The proposed system gets data from the Kinect that will be monitoring the hand and face gesture of the user while eating. Then, the gesture of hand and face data is sent to the microcontroller board to recognize and start counting bite taken by the user. The system recognizes the patterns of bite taken from user by following the algorithm of basic eating type either using hand or chopstick. This system can help people who are trying to follow a proper way to reduce overweight or eating disorders by monitoring their meal intake and controlling eating rate.

1. Introduction

The study of the bite counting is carried out due to escalating interest in this unique material and its immense potential in reducing weight. Obesity is the 5th major cause of death worldwide and about



2.8 million people die each year from a disease related to obesity [5]. Mostly all the diet plan approaches rely on calories from food label and this is not effective because not all food at grocery have label calories on package food. There is an ongoing debate of its recognition as a condition or a disease, motivated at least in part by the desire of researchers to increase options for its treatment and reduce the stigma and discrimination experienced by the obese [1]. It has been shown that monitoring and counting bite count reduce overweight people drastically. Our system bite counting consists of a simple algorithm which can detect hand gesture motion movement and facial movement and count as bite. The device can tell that target intake has been reached and it is time to stop eating, thereby helping people create long term healthy eating patterns and can prevent obesity. Z. Huang et al. [1] presented an approach which is based on body-worn sensors and mobile health technology. The component used was wrist motion tracker. N Salley et al. [2] presented an estimate of an individual's kilocalorie intake using bite count and mean kilocalories per bite determined by a formula based on demographic and physical characteristics using wrist motion tracker. While in L. Scisco et al. [3], they have proposed novel methods for measuring eating activity in free-living settings which also used wrist motion tracker to detect bite count while eating. Next, J. West et al [4] proposed this method to count bite with Participants that were recruited and asked to count and report their daily bites for one week. S. Sharma et al [6] use a wrist-worn device that can record 3-axis accelerometer and gyroscope. While, W. Accelerometers et al [7] they improve detection accuracy by combining using a head-mount accelerometer and using a wrist-worn accelerometer. S. Päßler et al [8] they use system of in-ear microphone that placed in the ear canal which can detect the sound of food intake Bedri et al. [9,10] called their system Outer ear interface that measure magnitude of deformation in ear canal during talk activities. The system able to classify all the classes up to 95% accurate. The system only recognizes jaw motion only.

2. Methodology

This study was divided into two steps: Facial jaw motion and hand gesture recognition motion using Microsoft Kinect 2.

2.1. Facial Jaw Motion

The Kinect sensor able to read and react to facial expression making it the most suitable sensor to use. The characteristic of jaw is capture using Kinect recognition ability because it can use infrared and don't need a light. Kinect can capture color component and even Skelton and joint of body. The jaw is located below the outer ear canal and Kinect detect a point in face. For example, if the face jaw is move when it will detect the movement of that point. Previous sensor use worn device which is not applicable in today's trend and for long term evaluation for daily used device.

2.2. Hand Gesture Recognition

Hand gesture system is widely use in today's trend since it only detecting hand, finger and gesture making it easier to identify the target. Depth image from Kinect sensor is used to detect the Skelton hand. The system worked well in dark area with no light making it advantage from other brand depth camera. The Kinect sensor can detect two hand gesture at current time. The Kinect video frame rate is 30Hz and above that can be used in real time and support high definition. The range to detect hand gesture is at good distance and easy to use.

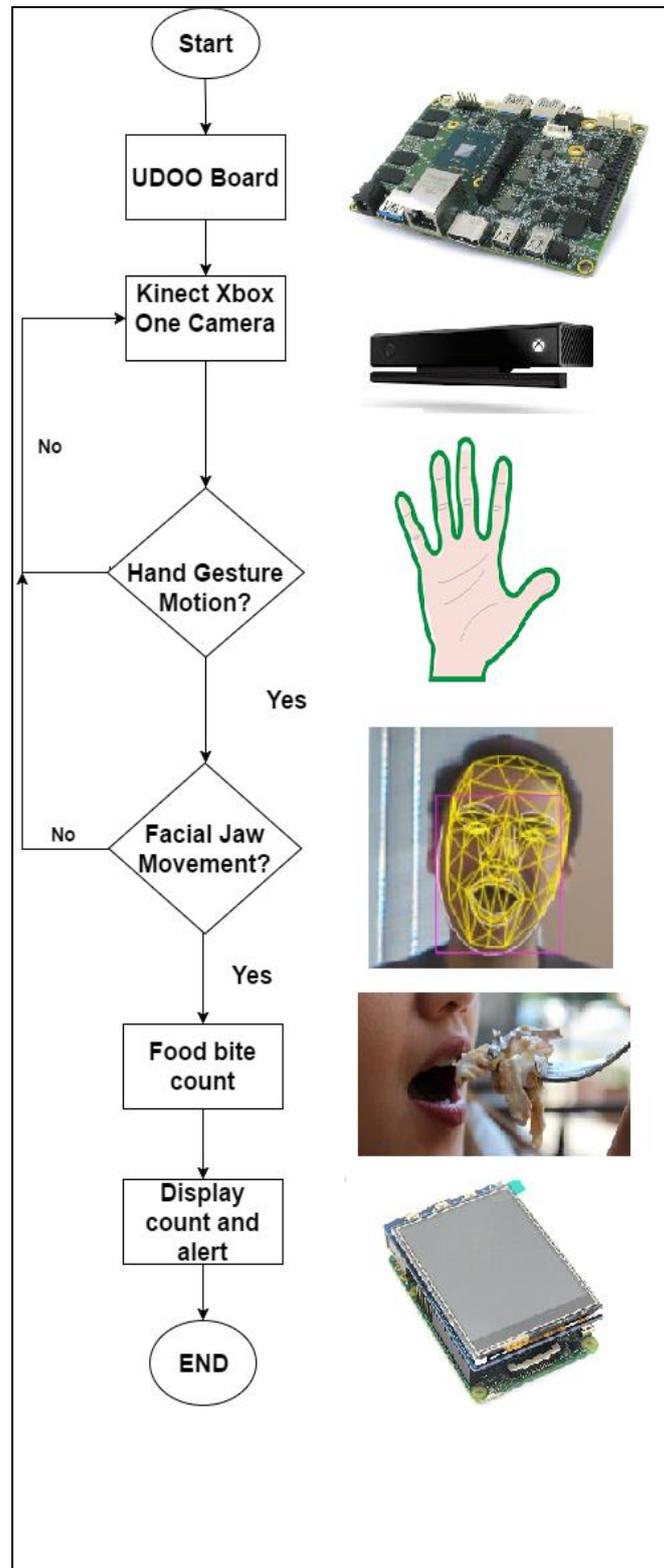


Figure 1: Flow chart of proposed project

The Kinect Xbox One sensor camera is connected and integrated into an embedded platform board. Our proposed project for this review Paper is shown in figure 1 above. The sensor will use depth

image to detect hand gesture while eating and facial jaw movement of person when person chew the food. The device is expected to detect the person while eating using monitoring system and will be able to alert the person of their bite count food intake. The depth images are captured from IR sensor and then stream to Kinect Xbox and can work at day and night without any problem. [18].

2.3. Comparison of project

Table 1 below show a comparison of type of sensor and classification of food intake among previous research. Currently there are five type of modalities approach such as sound, camera, inertial, fusion and depth sensing. Mostly are wearable type sensor attach to body except depth sensing approach which will be used in our proposed project. Depth sensing just only monitoring hand when eating and face jaw when user eat to detect the gesture and count it a bite. Fusion approach is a combined sensor in application. Inertial approach is a movement motion or gesture of body while camera approach is an image capture of food eaten to calculate the calories or to classify the volume of that food. The most used sensor approach was sound approach which divide into swallowing and chewing. Further detailed of previous research sensor type explained in topic wearable sensing approach.

Table 1. Overall comparison among modalities [43].

Type of modalities	Body Positions	Comfort	Applications	Current Applicability
Sound approach	Around Neck and ear (earpad, outer/inner ear canal)	Moderate, because need to wear during tracking.	For chewing and swallowing event	Still in laboratory testing and classification.
Camera approach	Attach in body or external by handphone device	High, because need to wear in shirt or need to hold the pone	Food classification and volume of food	Widely used in eButton device [16])
Inertial approach	Hand wrist, head and arm.	Moderate because need to attach with hand.	Food classification and bite counting	Still underestimate the capabilities of this method
Fusion approach	Flexibly (ear, wrist, external, etc.)	High because need to wear and attach in body	Chewing, swallowing and classification	Get good result tested in laboratory [17]
Depth sensing approach	Visual hand gesture and face gesture	No attach and friendly used	Bite event detection	No yet tested

2.4. Comparison of board

Table 2 provides an overall comparison among board with applicability of USB 3.0 feature and connectable to Kinect 2. The Udoo x86 comes with an embedded Arduino 101, with all upgrades enabled (including 6-axis accelerometer, gyroscope and Bluetooth). Connectivity functionalities are provided with an embedded Wi-Fi USB module, a Gigabit Ethernet port and a USB Bluetooth LE dongle. Using the connectivity of the Internet, it is possible to share and get data from the most common cloud services. It is based on Quad Core 64-bit new-generation x86 processors made by Intel®, designed for the PC domain. Prodigious processors concentrated in 14 nm, with an amount of energy consumption of 5 or 6 Watt. UDOO X86 can be used to program the Arduino™-compatible

module directly with the standard Arduino™ IDE. UDOO X86 mounts 8 GB eMMC on board, a Micro SD card reader as well as SATA, M.2 Key B and USB 3.0 ports to attach new-generation hard disk.

Table 2. Comparison of board.

Specification	Udoo x86	Lattepanda
CPU	Intel Pentium N3710 2.56 Ghz 4 core	1.8 GHz quad-core Intel Cherry Trail Atom X5 (64-bit)
RAM	2–8 GB	2–4 GB DDRL3L
USB Port	3x USB 3.0 port	2x (USB 2.0) and 1x (USB 3.0) ports
Network connectivity	Gb Ethernet + Bluetooth Low Energy + M.2 Key E slot for optional Wireless modules	Wi-Fi + Bluetooth 4.0
GPU	Intel HD Graphics 405	500MHz integrated GPU

3. Food Intake Mechanisms

This section is an overview on mechanism and classifying different type of sensor used by previous research. The research on wearable food intake usually classifying their sensor on 4 different kind of mechanical eating digestion process such as hand gesture, bite chewing and swallowing. The process is shown in figure 2 with the previous studies of different method of food intake classifying and in this paper the depth camera will be used for our research since it is not wearable and user friendly.

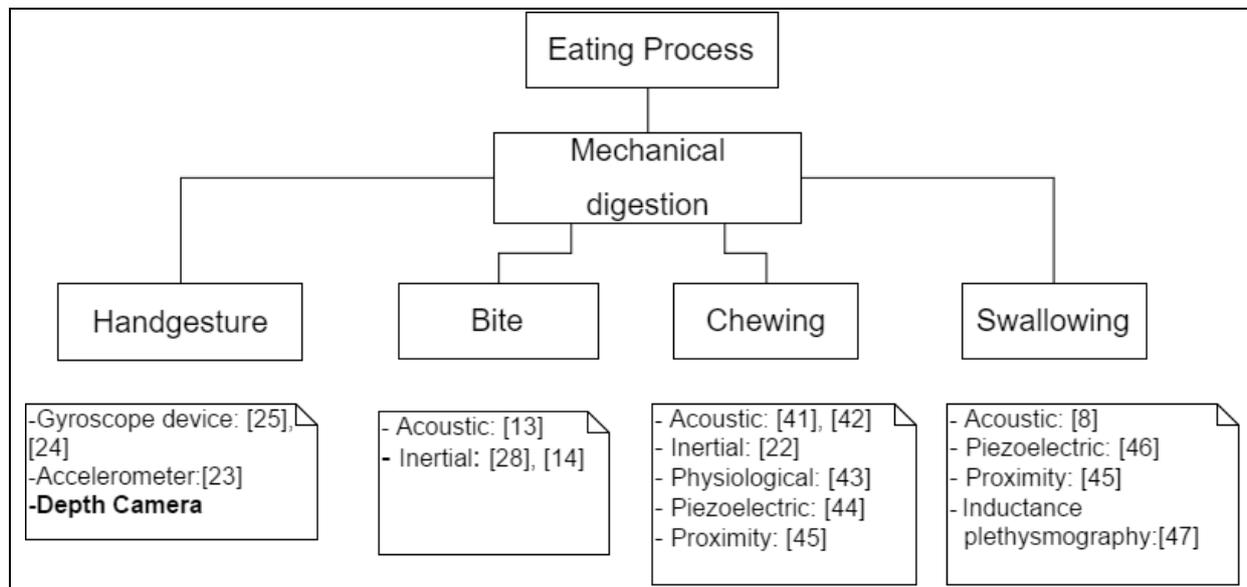


Figure 2: Classification of food digestion mechanism with previous trend and research studies.

4. Wearable Sensing Approaches

Section below describe a review of a different type sensor to detect food from previous research method and application

4.1. Acoustic/sound Approach

Chewing recognition: chewing and biting involving the movement motion of the jaw bone. Detecting the sound of chewing has a potential to the development of food intake monitoring. Mostly the research in this field use different algorithm such as microphone in ear and sound wave detection of bite to evaluated chewing event. Recently, O.Amft et al.[41,42] evaluated a chewing sound that use microphone attach in outer ear canal of user and they use many different algorithms on chewing sound detection. Next, O.Amft et al. [40] continue to demonstrate its recognition and classifying in food intake to simplifying individual reporting and they showed a good result in bite weight prediction on acoustic chewing recording. Moreover, Olubanjo et al. [39] focusing on noisy surrounding area such as in this paper restaurant background noise was implemented. The proposed method show a good result with 83.3% and 71.4% was achieved on a clean and noisy area. A total of 13 different food task were recorded in this research such as swallowing and chewing solid and liquid food. Next, Pabler et al. [8] proposed an automated chewing sound capture using microphone applied in outer ear canal. Other similar approach using method sound detection in food intake are proposed in [37,38].

Swallowing recognition: Swallowing involve of the three stages that is oral phase, pharyngeal phase and Oesophageal stage. During the oral phase, food is chewed and mixed with saliva to form a bolus. While, pharyngeal phase is where the vocal folds close to keep food and liquids from entering the airway and Oesophageal stage is where the bolus moves into the oesophagus. Many studies had developed a new way to detect food sound when swallowing. Sazonov et al. [36] proposed a method to capture the sound of swallowing by putting a microphone over the laryngopharynx and bites count are detect by placing below outer ear canal. Sazonov et al. [35] then proceed with a different method based on mathematical and they increase the accuracy from previous research. Bi et al. [34] proposed AutoDietary survey which have accuracy of 84.9% of food type recognition and a positive result on solid and liquid food detection.

4.2. Visual/Camera Approach

The camera does not focus on process but it is a method of observing a view of an images and video. Zhu et al. [32,33] created a dietary assessment using image segmentation that captured the meal from image using smart phone to estimate the food volume before and after meal taken. Energy of food is also determined with food classification. Zhu et al. [30,31] then explained the classification using Support Vector Machine(SVM) of food image using mobile device for dietary assessment. Shang et al. [29] proposed a system of Dietary Data Recording System (DDRS) that capture and calculated data of nutrient intake used when eating. They created mobile application in Android operating system that was user friendly and can be used in another smart phone. It also configures a 3D mesh reconstruction to measure and estimated the food intake using a video frame laser grid in smart phone camera. Other approach of camera based dietary was done by previous research in O'Loughlin et al. [28], Kikunaga et al. [27] and Wang et al. [26]. Mostly, they used camera for capturing image and analysis them for calculation to obtain the weight diet using selected method.

4.3. Inertial Approach

The development of inertial sensor was used in eating process to capture the movement such as motion and gesture in arm while eating. Sensor devices are basically used to determine the position and orientation of an object. Mostly sensors used to detect motion are the gyroscope and the accelerometer. Dong et al. [25] used a sensor of MEMS (microelectromechanical systems) which are better in detecting and tracking the wrist motion of hand when taking the food to mouth. The MEMS memory not large enough to store massive data however it can be used to track hand gesture and count the number of bites detected when user eating. On the other hand, Scisco et al. [24] come with alternative approach of a wrist-worn device and bite count engaging with device of motion sensor. The accelerometer was also part of inertial approach likewise, Thomaz et al. [23] the method used was a system of using 3-axis accelerometer built in digital watch which detecting the movement of a person

when eating. Mendi et al. [22] also used the accelerometer to capture the gesture of eating but then the data is sent to the smartphone via Bluetooth which is good because the device memory is limited to store big data.

4.4. Fusion Approach

The fusion approach was an action to merge a combined sensor in food intake monitoring. The purposes of combined sensor are to get rid the disadvantage of using one sensor method. The acoustic-visual approach was the clearest approach that can conclude from analysis. Moreover, Liu et al. [21] used a mix of sensor device with visual and sounds that were placed in outer ear for dietary habits of patient. Next, combine approach from Sen et al. [20] proposed the method of mix inertial sensor and image capture by phone. They used inertial sensor of the accelerometer and gyroscope for hand action and then the camera is automatically open for image evaluation.

5. Discussion

A lot variety of research in food tracking and monitoring. Recently, there is an ongoing research in food monitoring. knowledge into the tracking and counting of food is difficult because eating is affected by person, food, utensil, location and other factors, which could have been an action on the quality of any method used to measure the food intake. Actual scenario has a massive number of cuisine with different type and classified. It is a complicated process to track all the data from the sensor and each type of food for classification. Furthermore, the more the food type, the more burden the signal to process and will make the system less effective to handle. Mostly of the wearable sensors are limited to detect just calories of food intake, with the studies in [11,12]. There is still widely more item of technology can be used in this project such as depth sensor which can replace attach sensor of body. Most of the previous researches conducted on bite count detection is based on some sort of wearing devices such as worn body sensor, embedded sensors on body, neck, etc. The recent trend is now going towards non-wearable devices. Due to these reasons, bite detection systems based on vision or depth images are in high demand. One of the sensors that can capture depth images data and which can also track human skeleton is Kinect for windows [19].

6. Conclusion

This research involves detecting the bite count of hand motion during eating to control the over intake of food leading toward obesity. A bite-based measure of kilocalorie intake shows for individual use for self-monitoring to use for monitoring free-living kilocalorie intake. It is an easily collected and objective physiological signal based on wrist motion that could be refined to more accurately survey kilocalorie intake with the involvement of a measure of energy density and individual variables that are indicative of kilocalories per bite.

Acknowledgments

This research was financially sponsored by Universiti Tun Hussein Onn Malaysia (UTHM).

References

- [1] Z. Huang, "An Assessment of the Accuracy of an Automated Bite Counting Method in a Cafeteria Setting," 2013.
- [2] J. N. Salley, A. W. Hoover, M. L. Wilson, and E. R. Muth, "Comparison between Human and Bite-Based Methods of Estimating Caloric Intake," *J. Acad. Nutr. Diet.*, vol. 116, no. 10, pp. 1568–1577, 2016.
- [3] J. L. Scisco, E. R. Muth, and A. W. Hoover, "Examining the Utility of a Bite-Count-Based Measure of Eating Activity in Free-Living Human Beings," *J. Acad. Nutr. Diet.*, 2013.
- [4] J. West et al., "Pilot Test of A Bites-Focused Weight Loss Intervention," vol. 3, no. 4, pp. 1–4, 2015.
- [5] "Obesity and overweight", World Health Organization, 2017. [Online]. Available:

- <http://www.who.int/mediacentre/factsheets/fs311/en/>. [Accessed: 15- Feb- 2017].
- [6] S. Sharma, P. Jasper, E. Muth, and A. Hoover, “Automatic Detection of Periods of Eating using Wrist Motion Tracking,” pp. 5–6, 2016.
- [7] W. Accelerometers and G. Chen, “Automatic Eating Detection Using Head-Mount and,” pp. 578–581, 2015.
- [8] S. Päßler and W. J. Fischer, “Food intake monitoring: Automated chew event detection in chewing sounds,” *IEEE J. Biomed. Heal. Informatics*, vol. 18, no. 1, pp. 278–289, 2014.
- [9] A. Bedri, A. Verlekar, E. Thomaz, V. Avva, and T. Starner, “A Wearable System for Detecting Eating Activities with Proximity Sensors in the Outer Ear,” *Int. Symp. Wearable Comput.*, pp. 91–92, 2015.
- [10] A. Bedri, A. Verlekar, E. Thomaz, V. Avva, and T. Starner, “Detecting mastication: A wearable approach,” *Int. Conf. Multimodal Interact.*, pp. 247–250, 2015.
- [11] M. Sun et al., “eButton: A Wearable Computer for Health Monitoring and Personal Assistance,” in *Proceedings of the 51st Annual Design Automation Conference*, 2014, p. 16:1--16:6.
- [12] F. Kong and J. Tan, “DietCam: Automatic dietary assessment with mobile camera phones,” *Pervasive Mob. Comput.*, vol. 8, no. 1, pp. 147–163, 2012.
- [13] C. Hoebler et al., “Physical and chemical transformations of cereal food during oral digestion in human subjects,” *Br J Nutr*, vol. 80, no. 5, pp. 429–436, 1998.
- [14] Ye et al., “Assisting Food Journaling with Automatic Eating Detection,” *CHI Ext. Abstr. Hum. Factors Comput. Syst.*, pp. 3255–3262, 2016.
- [15] “Protein Digestion Monitoring Kit | Product”, Itsibio.com, 2017. [Online]. Available: <http://www.itsibio.com/prodm.html>. [Accessed: 06- Apr- 2017].
- [16] M. Sun et al., “eButton: A Wearable Computer for Health Monitoring and Personal Assistance,” in *Proceedings of the 51st Annual Design Automation Conference*, 2014, p. 16:1--16:6.
- [17] M. Farooq, S. Member, E. Sazonov, and S. Member, “Detection of Chewing from Piezoelectric Film Sensor Signals using Ensemble Classifiers,” pp. 4929–4932, 2016.
- [18] Y. Nizam, M. Norzali, H. Mohd, R. Tomari, and M. M. A. Jamil, “Development of Human Fall Detection System using Joint Height , Joint Velocity and Joint Position from Depth Maps,” vol. 8, no. 6, pp. 125–131, 1843.
- [19] Y. Nizam, M. Norzali, H. Mohd, and M. M. A. Jamil, “Classification of Human Fall from Activities of Daily Life using Joint Measurements,” vol. 8, no. 4, 1843.
- [20] S. Sen, V. Subbaraju, A. Misra, R. K. Balan, and Y. Lee, “The case for smartwatch-based diet monitoring,” *2015 IEEE Int. Conf. Pervasive Comput. Commun. Work. PerCom Work.* 2015, pp. 585–590, 2015.
- [21] J. Liu et al., “An Intelligent Food-Intake Monitoring System Using Wearable Sensors,” *2012 Ninth Int. Conf. Wearable Implant. Body Sens. Networks*, pp. 154–160, 2012.
- [22] E. Mendi, O. Ozyavuz, E. Pekesen, and C. Bayrak, “Food intake monitoring system for mobile devices,” *Proc. 2013 5th IEEE Int. Work. Adv. Sensors Interfaces, IWASI 2013*, pp. 31–33, 2013.
- [23] E. Thomaz, I. Essa, and G. D. Abowd, “A Practical Approach for Recognizing Eating Moments with Wrist-Mounted Inertial Sensing,” *Proc. 2015 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput. - UbiComp '15*, pp. 1029–1040, 2015.
- [24] J. L. Scisco, E. R. Muth, and A. W. Hoover, “Examining the utility of a bite-count-based measure of eating activity in free-living human beings,” *J. Acad. Nutr. Diet.*, vol. 114, no. 3, pp. 464–469, 2014.
- [25] Y. Dong, A. Hoover, J. Scisco, and E. Muth, “A New Method for Measuring Meal Intake in Humans via Automated Wrist Motion Tracking,” *Appl. Psychophysiol. Biofeedback*, vol. 37, no. 3, pp. 205–215, 2012.
- [26] D. H. (Okayama U. (Japan)) Wang, M. Kogashiwa, S. Ohta, and S. Kira, “Validity and reliability of a dietary assessment method: The application of a digital camera with a mobile phone card attachment,” *Journal of Nutritional Science and Vitaminology (Japan)*, vol. v. 48.

- [27] S. Kikunaga, T. Tin, G. Ishibashi, D.-H. Wang, and S. Kira, "The Application of a Handheld Personal Digital Assistant with Camera and Mobile Phone Card (Wellnavi) to the General Population in a Dietary Survey," *J. Nutr. Sci. Vitaminol. (Tokyo)*, vol. 53, no. 2, pp. 109–16, 2007.
- [28] G. O'Loughlin et al., "Using a wearable camera to increase the accuracy of dietary analysis," *Am. J. Prev. Med.*, vol. 44, no. 3, pp. 297–301, 2013.
- [29] J. Shang et al., "A pervasive Dietary Data Recording System," 2011 IEEE Int. Conf. Pervasive Comput. Commun. Work. PERCOM Work. 2011, pp. 307–309, 2011.
- [30] Fengqing Zhu, M. Bosch, Insoo Woo, SungYe Kim, C. Boushey, D. Ebert and E. Delp, "The Use of Mobile Devices in Aiding Dietary Assessment and Evaluation", 2017.
- [31] F. Zhu, M. Bosch, N. Khanna, C. J. Boushey, and E. J. Delp, "Multilevel segmentation for food classification in dietary assessment," 2011 7th Int. Symp. Image Signal Process. Anal., no. Ispa, pp. 337–342, 2011.
- [32] F. Zhu, M. Bosch, C. Boushey and E. Delp, "An image analysis system for dietary assessment and evaluation", 2017.
- [33] F. Zhu, M. Bosch, T. Schap, N. Khanna, D. Ebert, C. Boushey and E. Delp, "Segmentation assisted food classification for dietary assessment", 2017.
- [34] Y. Bi, M. Lv, C. Song, W. Xu, N. Guan, and W. Yi, "AutoDietary: A Wearable Acoustic Sensor System for Food Intake Recognition in Daily Life," *IEEE Sens. J.*, vol. 16, no. 3, pp. 806–816, 2016.
- [35] E. S. Sazonov, O. Makeyev, S. Schuckers, P. Lopez-Meyer, E. L. Melanson, and M. R. Neuman, "Automatic detection of swallowing events by acoustical means for applications of monitoring of ingestive behavior," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 3, pp. 626–633, 2010.
- [36] E. Sazonov, S. Schuckers, P. Lopez-Meyer, O. Makeyev, N. Sazonova, E. Melanson and M. Neuman, "Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior", 2017.
- [37] J. Nishimura and T. Kuroda, "Eating habits monitoring using wireless wearable in-ear microphone," 3rd Int. Symp. Wirel. Pervasive Comput. ISWPC 2008, Proc., pp. 130–132, 2008.
- [38] M. SHUZO et al., "Wearable Eating Habit Sensing System Using Internal Body Sound," *J. Adv. Mech. Des. Syst. Manuf.*, vol. 4, no. 1, pp. 158–166, 2010.
- [39] T. Olubanjo, E. Moore, and M. Ghovanloo, "Detecting food intake acoustic events in noisy recordings using template matching," 3rd IEEE EMBS Int. Conf. Biomed. Heal. Informatics, BHI 2016, pp. 388–391, 2016.
- [40] O. Amft, M. Kusserow, and G. Tröster, "Bite weight prediction from acoustic recognition of chewing," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 6, pp. 1663–1672, 2009..
- [41] O. Amft, "A wearable earpad sensor for chewing monitoring," *Proc. IEEE Sensors*, pp. 222–227, 2010.
- [42] O. Amft, M. Stäger, P. Lukowicz, and G. Tröster, "Analysis of Chewing Sounds for Dietary Monitoring," in *Proceedings of the 7th International Conference on Ubiquitous Computing*, 2005, pp. 56–72.
- [43] T. Vu, F. Lin, N. Alshurafa, and W. Xu, "Wearable Food Intake Monitoring Technologies: A Comprehensive Review," *Computers*, vol. 6, no. 1, p. 4, 2017.