DIET MONITORING THROUGH BREATHING SIGNAL ANALYSIS USING WEARABLE SENSORS

By

Bo Dong

A DISSERTATION

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

Electrical Engineering-Doctor of Philosophy

2014

ABSTRACT

DIET MONITORING THROUGH BREATHING SIGNAL ANALYSIS USING WEARABLE SENSORS

By

Bo Dong

This dissertation presents a framework of wearable food and drink intake monitoring system that analyzes human breathing signal for identifying swallows during the intake process. The system works based on a key observation that a person's otherwise continuous breathing cycles are interrupted by brief intra-cycle appears during the swallows. This dissertation develops wireless wearable electronics for capturing and processing human breathing signal, and algorithms for identifying intake-related swallows via recognizing apneas extracted from breathing signal. A family of apnea detection mechanisms including matched filters and machine learning has been developed. Algorithms are developed for detecting various types of swallowing events including for solid and liquid in the presence of many artifacts presents in free-living conditions. It is demonstrated that using these algorithms and the electronics, run-time intake monitoring and analysis are feasible at acceptable accuracy levels. Further accuracy improvements were explored using a Hidden Markov Model (HMM) based mechanism that leverages known temporal locality in the human swallow sequence. Finally, it was demonstrated that by combining swallowing signatures from breathing signal with hand movement signatures using accelerometers, it is possible to train a hierarchical Support Vector Machine (SVM) classifiers and a Hidden Markov Model (HMM) for accurate mealtime and duration estimation. The developed wearable system, along with a smartphone App, was experimentally validated on tens of subjects with approval from MSU's Institutional Review Board (IRB).

To My Family For All Their Love and Support

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Subir Biswas for his support and instruction through my PhD study. His academic rigor in research guided me to continue improving myself in terms of every aspect of research. I would also like to thank my committee Dr. Lalita Udpa, Dr. Sandeep Kulkarni, Dr. Nihar Mahapatra and Dr Rama Mukkamala for their time and support.

I would also like to thank my lab mates Qiong Huo, Mahmoud Taghizadeh, Debasmit Banerjee, Faeze Hajiaghajani, Stephan Lorenz, Muhannad Quwaider, William Tomlinson, Clifton Watson, Yan Shi, Saptarshi Das, and Dezhi Feng for participating in my experiments, brainstorming and implementation discussions.

Last but not least, I would like give my thanks to my friends for supporting me through this process.

TABLE OF CONTENTS

LIST OF T	ABLES	viii
LIST OF F	IGURES	. ix
Chapter 1: 1.1 1.2 1.3 1.4	Introduction Background Objectives Summary of Proposed Solutions Dissertation Structure	1 3 4 5
Chapter 2: 2.1 2.2 2.2.1 2.2.2 2.2.2 2.2.3 2.3	Related Works Invasive Swallow Detection Non-Invasive Swallow Detection using Chewing and Swallowing Sounds. Non-Invasive Swallow Detection using Other Modalities Non-Invasive Diet Monitoring using Infrastructural Sensors Proposed Mechanism	7 7 8 9 15 20 22
Chapter 3: 3.1 3.2 3.3 3.3.1 3.3.2 3.3.3 3.4 3.4.1 A. B. 3.4.2 3.4.3	Instrumentation System Breathing Process Swallowing Apnea Sensors for Collecting Respiratory Signal Conductive Rubber Sensor Piezoelectric PVDF Sensor Respiratory Inductance Plethysmography (RIP) Sensor Swallow Detection System Signal Shaping Circuit for Piezoelectric Belts Signal Shaping Circuit for Piezoelectric Belts Signal Shaping Circuit for RIP Belts µController and Bluetooth Module Breathing Signal Logging App	25 25 28 29 31 33 34 36 36 38 39 39
Chapter 4: 4.1 4.1.1 4.1.2 4.1.3 4.1.4 4.2 4.2.1	Swallow Detecting using Matched Filters Processing Methods Breathing, Apnea and Swallow Signature Matched Filter Method Machine Learning based Detection Artifacts Handling Results Results for Matched Filter based Detection	43 43 43 45 46 47 48 49

4.2.2	2.2 Performance for Machine Learning based Method with Time Domain Feat		
4.2.3	Performance for Machine Learning based Method with Frequency Domain	Features	
4.3	Discussion	52	
4.3.1	Iterative Template Refinement	52 52	
432	Discrimination Power of Time Domain Features	56	
4.3.3	Discrimination Power of Frequency Domain Features	58	
4.3.4	Artifacts Handling	60	
4.4	Summary	62	
Chapter 5:	Machine Learning Based Processing Algorithms	63	
5.1	Processing Methods	63	
5.1.1	Machine Learning Algorithms	63	
5.1.2	Breathing Apnea and Swallowing Signature	67	
5.1.3	Detection Scheme	69	
5.2	Experiments	70	
5.3	Results and Discussion	72	
5.3.1	Feature Extraction and Selection	72	
5.3.2	Swallow Detection	74	
5.4	Conclusion	76	
Chapter 6:	Support Vector Machine and Hidden Markov Model based Processing Al	gorithms 77	
6.1	Processing Methods	78	
6.1.1	Two-tier Swallow Detection	81	
6.1.2	SVM-based Swallow Detection with Posterior Probability	81	
6.1.3	Hidden Markov Model with Swallow Sequence Locality	84	
6.1.4	HMM Processing	88	
6.2	Results and Discussion	90	
6.2.1	Experimental Methods	90	
6.2.2	Performance Indices	91	
6.2.3	Feature Extraction for Stage-1 Detection using SVM	92	
6.2.4	Swallow Detection with SVM	95	
6.2.5	Improved Detection using HMM	98	
6.3	Conclusion	99	
Chapter 7:	Mealtime and Duration Monitoring1	01	
7.1	System Architecture 1	.02	
7.2	Processing Methods 1	.03	
7.3	Results	.09	
7.3.1	Experimental Methods 1	.09	
7.3.2	Performance Evaluation 1	11	
7.3.3	Feature Extraction 1	13	
7.3.4	Performance of Food Intake Detection 1	14	
7.3.5	Performance of Meal Intake Analysis 1	18	

7.4	Discussion	
7.4.1	Restrictive Feature Selection	
7.4.2	Benefits of Hierarchical Classifier	
7.4.3	Performance of Existing Research	
7.4.4	Spontaneous Swallows	
7.5	Conclusion	
Chapter 8:	Proposed Work	
8.1	Diet Volume Detection	
8.2	Choking and Coughing Detection	
BIBLIOGR	APHY	

LIST OF TABLES

Table 4-1: Performance of classifiers using time domain features
Table 4-2: Performance of all three classifiers using frequency domain features. Performance on both individual dataset and the combined dataset are presented in this table
Table 5-1: Durations of different breathing cycle types 71
Table 5-2: Features selected for classification
Table 5-3: Performance of the first stage of the hierarchical classifier
Table 5-4: Performance of the second stage of the hierarchical classifier
Table 6-1: Comparison between fixed threshold SVM-only and two-tier SVM+HMM mechanism 98
Table 7-1: Features Extracted For Svm Classifiers 113
Table 7-2: Comparison Between Svm-Only And Svm+Hmm Solutions 116
Table 7-3: Comparison of 3-Stage Hierarchical Classifier And Single Classifier

LIST OF FIGURES

Figure 1-1: Prevalence of obesity worldwide (From International Association For the Study of Obesity)
Figure 1-2: Increase of obesity in US
Figure 1-3: Dissertation flowchart
Figure 2-1: Schematic sensor positioning in [29] 10
Figure 2-2: In-ear microphone used in [33]13
Figure 2-3: sensors used in [36]14
Figure 2-4: Sensor positioning in [39] 16
Figure 2-5: Prototype gyroscope sensor used in [42] 19
Figure 2-6: Magnetic coils and microphone system proposed in [46] 20
Figure 2-7: Camera based system proposed in [47] 21
Figure 2-8: The embedded RFID and weighing table surfaces
Figure 3-1: 3 steps of swallowing
Figure 3-2: Breathing signal of two subjects in the experiments
Figure 3-3: Resistive stretch sensor (a) and its static response (b)
Figure 3-4: Transient response of stretch sensors
Figure 3-5: Piezoelectric belt (a) and its static response (b)
Figure 3-6: Equivalent circuit of piezoelectric sensor (a), Isolation circuit (b), and Transient response of piezo respiratory sensor(c)
Figure 3-7: RIP sensor
Figure 3-8: System architecture of swallow detection system using piezoelectric belt
Figure 3-9: System architecture of swallow detection system using RIP belts
Figure 3-10: Signal shaping circuit for piezoelectric belt
Figure 3-11: Schematic of signal shaping circuit

Figure 3-12: Logic layering of LiveActive	40
Figure 3-13: Design of the webserve module	41
Figure 3-14: Graphic User Interface (GUI) of LiveActive	41
Figure 4-1: Examples of Breathing Cycles (BC), Normal Breathing Cycles (NBC), Breathing Cycles with Inhale Swallow (BC-IS), Breathing Cycles with Exhale Swallow (BC-ES) a apnea	ing ind 44
Figure 4-2: Detection process of matched filter based detection algorithm	46
Figure 4-3: Breathing signal variability before, during, and after talking	48
Figure 4-4: ROC Distribution for all seven subjects with arbitrary templates.	50
Figure 4-5: Similarity score space for: a) initial matched filter template used as a starting poi and b) the final template obtained at stabilization of the iterative algorithm. The tighter cluster of the points in the bottom graph indicates iterative improvement of the template quality	int, ing 55
Figure 4-6: Iterative template refinement performance; a) true positive rate, and b) false positive rate evolution with iterations.	ive 56
Figure 4-7: Utility of the time domain features for Subject-1; three peaks in the left figure a caused by different types of breathing cycles with feature distribution shown in the right figure.	are ire. 57
Figure 4-8: (a) utility of frequency domain features, (b) comparison between time and frequend domain features; results are presented with limited number of features that are chosen using method as described.	icy g a 58
Figure 4-9: Power spectral density (PSD) of breathing signals with talking and without talking when normal breathing or breathing with swallows are executed.	ng, 60
Figure 4-10: Breathing signal: (a) with upper body rocking movement, and (b) without movement.	the 61
Figure 5-1: Examples of Breathing Cycles (BC), Normal Breathing Cycles (NBC), Breathing Cycles with Inhale Swallow (BC-IS), Breathing Cycles with Exhale Swallow (BC-ES) a apnea	ing ind 67
Figure 5-2: Example breathing signals for solid and liquid swallows	68
Figure 5-3: Logic for swallow signature detection	70
Figure 5-4: Discriminative property of time and frequency domain features	72
Figure 5-5: Benefits of ±1 crossings as a classification feature	73

Figure 6-1: Respiratory signal with swallow signature
Figure 6-2: Processing scheme for swallow detection
Figure 6-3: (a) Hidden breathing state machine and (b) HMM processing components
Figure 6-4: Experimental setup
Figure 6-5: Feature discriminative property and ± 10 crossings as a classification feature
Figure 6-6: Distribution of posterior probabilities with and without swallows
Figure 6-7: Comparison between SVM-only and two-tier SVM+HMM mechanism
Figure 7-1: Components of the mealtime and duration monitoring system 103
Figure 7-2: The mealtime and duration detection scheme
Figure 7-3: Meal intake analysis algorithm
Figure 7-4: Experimental setup 110
Figure 7-5:. Performance of SVM-only food intake detection method
Figure 7-6: An example temporal dynamics of the meal intake analysis process
Figure 7-7: Threshold selection for different window sizes
Figure 7-8: Performance of meal intake analysis module
Figure 7-9: Performance of Classifier-1, 2 and 3 with different feature count
Figure 8-1: Anticipated output of chest belts during coughing (a) one coughing within a breathing cycle (b) two consecutive coughs in a breathing cycle

Chapter 1: Introduction

1.1 Background

According to the data from World Health Organization, worldwide obesity increased over 200% since 1980 [1]. Figure 1-1 shows the prevalence of obesity in different countries based on the data collected by International Association for the Study of Obesity. Observe that obesity is currently less prevalent in Asian countries such as China, India, Japan etc, while it is a severe problem in countries such as United States, Canada, Australia, etc. In particular, in 2010, 35.5% of men and 35.8% of women in US are obese.



Figure 1-1: Prevalence of obesity worldwide (From International Association For the

Study of Obesity)

The prevalence of obesity in US is constantly increasing. Figure 1-2 depicts the increase in percentage of obesity in population from 1960 to 2010. Similar trends are also predicted for developing countries, such as China [2][3] and India [4][5].



Figure 1-2: Increase of obesity in US

The prevalence of obesity brings many health problems, both physical and mental, and social issues. Eckel *et al* [6] had proven that obesity is a major risk factor for coronary heart disease, and it is known that 5% to 10% weight reduction can decrease blood pressure and total blood cholesterol, improve glucose tolerance for patients with diabetes, and reduce the severity of obstructive sleep apnea. Visscher [7] mentioned that obesity can cause cardiovascular disease, type-2 diabetes mellitus, cancer, osteoarthritis, work disability and sleep apnea, and it has pronounced impact on morbidity. Moreover, Jia [8] has shown that obesity also affects health-related quality-of-life (HRQL) [9], which is a multi-dimensional concept that includes domains related to physical, mental, emotional and social functioning, and assesses the positive aspects of a person's life, such as positive emotions and life satisfaction. It is proved that people with

obesity had significant lower HRQL than those who had normal weight, and such low HRQL were also seen for people without chronic diseases caused by obesity.

There are many factors found to be associated with obesity. Nielsen *et al* [10] demonstrated that short sleep duration was consistently associated with development of obesity in children and young adults. It has also been found out that environmental, perinatal and genetic factors induce neuroendocrine perturbation followed by abdominal obesity [11]. However, Astrup *et al* [12] pointed out that the prevalence of obesity we were facing should not be caused by genetic factors, because the development of obesity worldwide is too rapid to be associated with genetic changes, and there were only few humans had been shown to have genetic obesity. From an energy balance point of view, obesity is caused by the imbalance between the energy we derive from food and drink and the energy we expend for metabolism and physical activity [12][13].

1.2 Objectives

Diet control and physical exercise are the two most important components of obesity control. Traditionally, self-reported questionnaires were widely used by researchers for estimating both food intake and physical activity levels for high-risk individuals. In most such studies, participants have shown underreporting tendency. Additionally, self-reporting by elderly population is often unreliable due to poor memory situations. These make questionnaire-only based methods subjective and unreliable [14][15][16].

In recent years, accelerometers, gyroscope, and pressure sensors have been widely utilized for instrumented physical activity monitoring with high detection accuracy [17][18][19][20]. However, not many efforts on instrumented diet monitoring are reported in the literature. Diet monitoring can reduce the subjectivity [14] associated with questionnaire based self-reporting systems.

An instrumented system can potentially detect each instance of food/drink intake, and can have significant impact on obesity and overall health monitoring and management. Together with self-reporting of dietary habits at a high level, the system can quantify calorie intake trends and estimates for its users. It has also been proven that such health monitoring can improve the effectiveness and quality of healthcare service [21].

We present a wearable sensor system for solid food intake monitoring based on swallows detected in breathing signals. Using a wearable chest-belt, we detect swallows by the way of detecting apneas extracted from breathing signal captured by the chest-belt. Since the belt can be worn inside, outside, or between garments (it does not need skin contact), it has the potential for prolonged comfortable daily usage without raising any cosmetic and discomfort concerns.

1.3 Summary of Proposed Solutions

The sensor system and intake monitoring algorithms developed in this thesis works based on a key observation that a person's otherwise continuous breathing process is interrupted by a short apnea during a swallow, which is a part of the swallowing process [22]. We first detect swallows by the way of detecting apneas extracted from breathing signal captured by a wearable wireless chest-belt. Afterwards, swallow pattern analysis is used for identifying solid/liquid swallows. Together with self-reporting at the high level of overall diet habits (i.e., the types of food and drinks etc.), the instrumented detection of swallow counts can offer an objective way to: 1) study the food and drink intake trends, and 2) estimate calorie intake. In this thesis, however, we only address the automatic swallow detection part of the process.

1.4 Dissertation Structure

Figure 1-3 illustrates the structure of the thesis:



Figure 1-3: Dissertation flowchart

Chapter 1 introduces the prevalence of obesity and overweight, the causes, and the drawbacks of existing questionnaire based methods.

Chapter 2 first investigates the existing invasive and non-invasive swallow detection methods, and analyzed their drawbacks. Then it talks about the proposed swallow detection mechanism, and its advantages over the existing methods.

Chapter 3 first introduces the physiological process of breathing and swallowing, and brings up the concept of swallow apnea. It then depicts the sensors and swallow detection system for diet monitoring.

Chapter 4 introduces the matched filter method used for liquid swallow detection. Machine learning method is also proposed for comparison. Iterative template refinement, feature selection, and artifacts handling are also discussed in this chapter.

Chapter 5 proposes a hierarchical classification algorithm for solid and liquid swallow detection. It compares the performance of different features and different popular machine learning algorithms.

Chapter 6 provides the method of cascading Support Vector Machine (SVM) with Hidden Markov Model (HMM) for improving the accuracy of food intake detection. The modeling, processing, and performance of this method are discussed in details.

Chapter 7 proposes a hierarchical classification method cascaded with HMM, diet time and duration analysis are performed to indicate the dietary habits. The algorithms are validated through least controlled experiments. Artifacts such as spontaneous swallows, talking, laughing, coughing and clearing throat are considered in the algorithm.

Chapter 8 summarizes the thesis and discusses the future work that can be done.

Chapter 2: Related Works

As mentioned in Chapter 1, diet control and physical exercise are the two most important aspects in obesity control. With the development of microelectromechanical systems (MEMS), accelerometers have been widely used in physical activity detection and energy expenditure estimation. Zhang *et al* [23] had shown that together with some other sensors (such as temperature sensors, heart rate sensors) the IDEEA[®] (Intelligent Device for Energy Expenditure and Physical Activity) device is able to provide over 95% accuracy in energy expenditure estimation. In this disseration, we are focusing on the swallow detection, which is strongly correlated with food and drink intake. Existing swallow detection methods are generally classified into two groups: invasive and non-invasive.

2.1 Invasive Swallow Detection

Many patients with neurologic issues due to stroke, multiple sclerosis, trauma, bulbar palsy and other impairment may have difficulties in swallowing. Videofluoroscopy is therefore used in [24] to provide important information on impairment of the swallowing process, providing essential information to the doctors for arranging treatment accordingly. The paper (a) described the indications for videofluoroscopic swallowing studies by evaluating patients with neurological disorders affecting swallowing, (b) described the techniques for evaluating the swallow mechanism with videofluoroscopy in a standardized manner, and (c) used cine videofluoroscopy to illustrate the range of abnormalities that can be demonstrated for some of these conditions and discussed the effect of patient treatment. This method provides the images to demonstrate each stage of swallowing, and shows the movement of bolus in detail. However, it requires swallowing food or water mixed with barium, which labels the bolus under X-ray, and patients need to be exposed under radiation. For those reasons, this method is not applicable for swallow detection for obesity control purposes. Perlman et al [25] analyzed the duration and temporal relationship of electromyographic activity from the submental complex, superior pharyngeal constrictor, cricopharyngeus, thyroarytenoid and interarytenoid muscles during swallowing saliva, 5ml and 10ml water. Bipolar, hooked-wire electrodes were inserted into the muscles mentioned above except the submental complex, which was analyzed with bipolar surface electrodes. The experiment included 8 healthy subjects executing 5 swallows for saliva, 5ml and 10ml water individually composing a 120-swallow data set. The total activation duration of all the muscles during pharyngeal phase of the swallow did not change with bolus size, but some muscles did demonstrate a difference in electromyograph duration and time of firing between saliva and 10ml water. Submental muscle activity was longest for saliva swallows. The interarytenoid muscle showed a significant difference in duration between the saliva and 10ml water swallow. Finally, the interval between the start of laryngeal muscle activity and pharyngeal muscle firing pattern decreased as the bolus volume increased. The muscle activation pattern showed a high correlation within dataset of each subject and high variance across different subjects.

2.2 Non-Invasive Swallow Detection

Non-invasive methods use accelerometers, microphones, surface electromyography, piezoelectric sensors etc, to collect physiological signals related to swallowing without involving tools that break the skin or physically enter the body. Comparing to invasive methods, the non-invasive alternatives have some important advantages:

Safety: invasive methods needs tools such as electrode needle to penetrate the skin to collect electromyography signal, or patients need to take barium to label bolus under X-ray,

8

which should be carried out by medical providers with special trainings. Incorrect operations, such as contaminated electrode needles and long-term exposure to X-ray, may severely impact health of the patient even cause death. However, non-invasive methods do not need to break the skin, and misplacement of the devices will not lead to negative impacts.

Ease of use: experiments using invasive methods need to be executed in hospitals under the supervision of professionals, while the non-invasive devices can be even embedded into clothes [26] or packaged as hearing aid devices [27]. Swallowing detection algorithms can be implemented in microcontrollers, and results may be wirelessly transferred to health care providers for analysis.

Long-term monitoring: as invasive experiments need to be carried out in health care facilities with the help of specially trained professionals, they are not suitable for long-term monitoring. However, non-invasive sensors can be designed to be portable and suitable for long-term usage.

Non-invasive sensors also have the drawback of not being able to provide as much information as invasive methods do, i.e., non-invasive methods cannot provide detailed information about the activation of muscles involved in swallowing process. But non-invasive sensors provide a suitable solution for the application of food intake monitoring for diet control, which mainly focus on the number and timing of swallowing events.

2.2.1 Non-Invasive Swallow Detection using Chewing and Swallowing Sounds

Takahashi *et al* [28] suggested that cervical auscultation in the evaluation of the pharyngeal swallow may become a part of the clinical evaluation of dysphagic patient. The presented study investigated three aspects of the methodology for detecting swallowing sounds: (1) the type of acoustic detector for the analysis of pharyngeal swallow, (2) the type of adhesive

suitable for attaching the sensor, (3) the optimal sites for detection. An accelerometer with double-side paper tape was selected as the optimal detector due to its wide range of frequency response and small attenuation level. Using this sensor, swallowing sounds and noises were collected at 24 sites on the neck for 14 healthy subjects. The optimal position for collecting swallowing sounds were selected with the largest signal to noise ratio and the smallest variance, and it was proven that the site over the literal border of the trachea immediately inferior to the cricoids cartilage is the optimal site. The site over the center of the cricoid cartilage and the midpoint between the site over the center of the cricoid cartilage and the site immediately superior to the jugular notch were also considered to be the most appropriate sites. This method provides some guidance to the study of swallowing detection using acoustic information, but the sensors are needed to be attached in the neck region, which brings cosmetic and safety issues.



Figure 2-1: Schematic sensor positioning in [29]

Cichero *et al* [30] presented a hypothesis of the cause of swallowing sounds. It was suggested that as pharynx contains a number of valves and pumps that produce vibrations and

reverberations within the pharynx to generate swallowing sounds. An analogy was proposed between swallowing sounds and heart sounds that propagate via vibration of muscles and valves. Therefore, many literatures use swallowing sounds as the metric of detection food intake. In order to derive clear swallowing sounds, microphones are normally attached to the neck region to be close to pharynx.

Amft et al [29] presents an investigation to detect and classify normal swallowing during eating and drinking from surface electromyography (SEMG) and microphone sensors as shown in Figure 2-1. Gel electrodes were placed in the submental and infra-hyoid regions, and recorded at 24 bit 2kHz. Swallowing sound was recorded using an electrets condenser microphone placed inferior midline from the cricoid cartilage. The non-invasive sensors were selected to be integrated into a collar-like fabric for continuous monitoring of swallowing activity over long period. Feature similarity detection mechanism was applied for both SEMG signal and sound signal. Signals were first segmented using Sliding-Window And Bottom-up (SWAB) algorithm [31], which partitioned continuous stream of sensor data by sequentially testing the approximation of the signal through linear regression lines and using the boundaries of these approximations as segments. Each segment of sensor data was then compared to a trained pattern using Euclidean distance for calculating feature similarities. The detection results based on SEMG and sound signals are then fused using parameter training method. Overall 80% recall and 70% precision were achieved using the proposed methods. This paper further utilized machine learning algorithms to differentiate high vs. low volume swallows, and high vs. low viscosity bolus with normalized accuracy around 70%. This paper is one of the earliest works on swallow detection. However, as the electrode needs gel to guarantee the electric contact with the skin and the microphone needs to be closely attached to the neck, the system may not be

comfortable for prolonged usage. Moreover, because the sensor system is worn on the neck region, it will bring up some cosmetic issues.

Passler et al [32] proposed a method for non-invasive monitoring of human food intake behavior and long-term dietary protocol by using only chewing and swallowing sound sensors. The sensor system was belt using an in-ear microphone and a reference microphone integrated in a hearing aid case for recording chewing and swallowing sounds in the ear canal and environmental noise. It was observed that food intake sounds recorded by the in-ear microphone had slightly higher signal amplitude than the same sounds recorded by the reference microphone, while environmental sounds and speech of the participant have comparable signal energies in both records. Another parameter, magnitude squared coherence function (MSC), was also used, as the MSC of the environmental sound is high, while that of the food intake sound is low. It was demonstrated that the detection algorithm by comparing the signal energy of the two microphones outperformed the method using MSC, and the precision and recall can be 91.3% and 81.8% respectively. However, this method suffers from some drawbacks: (1) the experiment was done in quiet room or office without other disturbing environmental noise; (2) the performance of the system on differentiating swallowing sounds and chewing sounds was not illustrated. It is very common that some people chew food for longer time, while some other people swallow before food is fully crunched.

Similarly, Nishimura *et al* [33] developed an in-ear microphone embedded into a common Bluetooth headset which is used to capture sound emission generated by chewing as shown in Figure 2-2. A two stage recognition algorithm was proposed. First, the "Chew-like" signal detection was performed by using the number of zero-crossings of the regression coefficients with a negative valued slope and the local peak of LPF output. Second, the chewing sound verification was performed by comparing the extracted features of the testing sound signal with the training sound signal. High accuracy of 98.7% was reported for 5 food categories, including chips, salad, rice, wafers, and banana. However, details about the verification process were not provided. The author did not mention the details of the experiments either, such as the number of subjects, and how the training data set and testing data set were formed.



Figure 2-2: In-ear microphone used in [33]

Makeyev *et al* [34] presented a fully automatic food intake detection methodology, with the aim of improving understanding of eating behaviors associated with obesity and eating disorders. The system proposed used a miniature throat microphone attached over the laryngopharynx, which had a dynamic range of 46±3dB with a frequency range of 20-8000Hz. The proposed method consisted of two stages. First, acoustic detection of swallowing instances based on mel-scale Fourier spectrum features and classification using support vector machines was performed. Principal component analysis (PCA) and smoothing algorithms were performed to improve detection accuracy. Second, the frequency of swallowing is used as a predictor for detection of food intake interval. Experiments were carried out on 12 subjects with different degree of adiposity. Average accuracy of >80% and 75% were obtained for intra-subject and inter-subject models. However, similar to [29], the microphone attached to the neck may not be comfortable for long-term usage, and it also brings up cosmetic issues. External noise was not considered either during the experiments.

Walker *et al* [35] also used the throat microphone system for collecting swallowing sounds. Short Time Fourier Transform (STFT) was performed on the collected audio data, and it was found that swallow sounds have a stronger presence in the upper frequencies comparing with other sounds such as vocal cord activation (hum, whispering, and speaking), cleaning the throat, and coughing. Discrete Wavelet Transform (DWT) was then used to obtain higher temporal resolution than that offered by STFT at high frequency intervals. Windowed signal energy and windowed maximum are used to perform swallow event detection. The proposed mechanisms were tested on two male subjects. The main drawback of this work is that only two subjects were included in the experiments, and environmental noises are not fully considered.



Figure 2-3: sensors used in [36]

Sazonov *et al* [36] developed a swallowing and chewing monitoring system to study the behavioral patterns of food consumption and producing volumetric and weight estimates of energy intake as shown in Figure 2-3. The system worked based on detecting swallowing by a sound sensor located over laryngopharynx or by a bone conduction microphone and chewing through a below-the-ear strain sensor. The system can be implemented in a wearable monitoring

device, thus suitable for monitoring ingestive behavior in free living settings. Experiments were carried out on 21 subjects during eating and quiet sitting. Video and sensor data were manually labeled by trained professionals. The reliability of manual labels was tested on 5 subjects and it was demonstrated that the intra-class correlation coefficients are 0.996 for bites, 0.988 for chews and 0.98 for swallows. The collected sensor signals and the resulting manual scores were left for future research.

Aboofazeli *et al* [37] presented a Hidden Markov Model (HMM) based method for the swallowing sound segmentation and classification method. Swallowing sounds of 15 healthy and 11 dysphagic subjects were studied. The swallowing sound signals were segmented into 25 ms segments, and 7 features were extracted. Trained HMM model classified the sound signals into three phases: initial quite phase, initial discrete sound (IDS) and bolus transit sound (BTS). Multi-scale products of wavelet coefficients were proved to be the most effective feature for HMM. HMM model was also built to differentiate the swallowing sound of healthy subjects and patients with disphagia, and accuracy of 85.5% was achieved. However, as the experiment was done in strictly controlled environment, and artifacts such as ambient noise, talking were not included, the method may not be used for everyday diet monitoring.

2.2.2 Non-Invasive Swallow Detection using Other Modalities

Other than sounds generated during chewing and swallowing, inertial sensors were also widely used to monitor other physiological phenomenon related to eating and drinking.

Amft *et al* [38] proposed a two-stage recognition system for detecting food intake related arm gestures. Information derived from this system can be used for automatic food intake monitoring in the domain of behavioral medicine. It is demonstrated that arm gestures can be clustered and detected using inertial sensors on the arm. Experiments were carried out on 2 subjects with 384 gestures with 4 sensors attached to the right and left lower and upper arm. The subjects were asked to eat or drink using cutlery, spoon, hand and glass. An accuracy of 94% can be achieved by using HMM method. When analyzing the continuous data, an accuracy of 87% can be reached. However, the experiments were done on a limited data set, i.e., only two subjects were included in the experiments, and the type of food lacks variety. Moreover, the experiment was done in controlled environment, and there would be many false positives if gestures such as smoking, scratching the head or face, etc were analyzed.



Figure 2-4: Sensor positioning in [39]

Amft and Troster [39] proposed a dietary monitoring system using more than one modalities as shown in Figure 2-4. The on-body sensing approach was chosen based on three core activities during food intake: arm movement, chewing and swallowing. The arm and trunk movements associated with food intake were measured using inertial sensors, i.e., accelerometer, gyroscope and magnetometer, chewing sounds were recorded using an in-ear microphone, and swallowing activities were acquired by a sensor-collar containing surface electromyography

(SEMG) electrodes and a stethoscope microphone. In three independent evaluation studies, the continuous recognition of activity events had been investigated and performances were evaluated. An event recognition procedure was deployed that addresses multiple challenges of continuous activity recognition, including the dynamic adaptability for variable-length activities and flexible deployment by supporting one to many independent classes. The approach uses a sensitive activity event search followed by a selective refinement of the detection using different information fusion scheme. With experiments, four intake gesture categories from arm movments and two food types from chewing sounds were detected with a recall of 80-90%, and a precision of 50-64%. 68% of recall and 20% of precision was achieved for individual swallows. Although this work is one of the most comprehensive work in dietary monitoring to our best knowledge, it suffers from the following drawbacks: (1) the system used three groups of sensors on the arms, neck and in the ear respectively, which may not be convenient for everyday usage and cause cosmetic issues, (2) the proposed algorithms had low precision for swallow detection, meaning many false positives would be expected.

Mioch *et al* [40] examined the patter of activity in masseter and temporalis muscles during mastication of different food samples with known textural properties and analyzed the interindividual variations. Surface electromyography (SEMG) signals were recorded from the right and left masseter and temporalis muscles in 36 young adults during free-living and side-imposed mastication. 5 different types of food with known rheological properties were used. Both masseter and temporalis activity increased with increased stress at measurements of food, which confirmed that the mastication process was adjusted according to the food texture. Temporalis muscle activity was more influenced by food texture than masseter muscles. Less muscle activity was observed to chew the food during free-living scenario. However, 25% of the subjects did not show any differences between side-imposed mastication and free-living scenarios, indicating that they may have greater chewing efficiency on one side. Therefore, measuring the activities of masseter and temporalis muscles may be used for analyzing the food intake of subjects, however, there are cases that foods with the same texture have very different calorie densities, such as mushroom and meat.

Nahrstaedt *et al* [41] investigated the use of a combined electromyography (EMG) and bioimpedance (BI) measurement at the throat to automatically detect swallowing events. The measured BI indicated the closure of larynx. There is a typical drop in BI during swallowing. The activations of the muscles involved were measured using EMG. Valley detection algorithm was used to segment BI signals. Additionally, only BI valleys that coincide with EMG activations are selected for feature extraction. Then the extracted features from BI and EMG signal were classified using Support Vector Machine (SVM) to identify BI valleys related to swallowing events. The proposed methods were tested on 9 healthy subjects. The data set contained 1370 swallow events with different bolus size with artifacts such as movement and speech. The combined BI/EMG segmentation detected 99.3% of all swallow events. The subsequent SVM classifier had a sensitivity of 96.1% and a specificity of 97.1%. However, similar to other methods based on EMG, skin contact using conductive gel is required and the suitability for long-term usage is questionable.

Dong *et al* [42] proposed a method for measuring food or drink intake through automated tracking of wrist motion as shown in Figure 2-5. A watch-like device with microelectromechanical (MEMS) gyroscope was used to detect and record the motion of hand, which was believed to be related to food or drink intake. This method was found to have 94% sensitivity in controlled meal setting and 86% sensitivity in uncontrolled setting, and both had

one false positive out of every 5 bites. Preliminary data showed that bites measured by the device were positively related to calorie intake indicating the potential of the device to monitor energy intake. However, the watch was designed to be on during eating only.



Figure 2-5: Prototype gyroscope sensor used in [42]

Moreau-Gaudrey *et al* [43] proposed a user-friendly non-invasive bedside procedure for studying swallowing and swallowing disorders in the elderly considering the frailty of this age group. In this study, respiratory inductance plethysmography (RIP) was proposed. An automated process for the detection of swallowing was designed, and the first derivative of the breathing signal was used to pick up the apnea during breathing. An accuracy of 90% was reported given that an appropriate threshold had been selected. However, only 56 swallows from 14 subjects were recorded, and no artifacts, such as motion and speech, were considered. Moreover, people breathing at a lower rate may have longer apneas between consecutive breathing cycles, which can be detected as false positives using the proposed mechanism.

Damouras *et al* [44] proposed a method of using an accelerometer for swallowing detection. A dual-axis accelerometer was attached to the participant's neck (anterior to the cricoids cartilage) using double-sided tape. In the paper, the acceleration signal was considered as a stochastic diffusion where movement was associated with drift and swallowing with volatility. Consequently, a volatility-based online swallow event detector that operated on the raw acceleration signal was developed. With data from healthy subjects and subjects with

dysphagia, the proposed method is proved to be working as good as their previous work [45], where same data was used, but without preprocessing. However, the experiment was done in a more strictly controlled environment, and less artifacts were considered.



Figure 2-6: Magnetic coils and microphone system proposed in [46]

Kandori *et al* [46] developed a swallowing detection system that can detect swallowing sounds and measure the distance between two magnetic coils as shown in Figure 2-6. The coils were set on both sides of the thyroid cartilage, and the distance between them changes in accordance with the movement of the thyroid cartilage. Swallowing sounds were detected by a piezoelectric microphone attached to the neck. The coils and microphones were installed on a holding structure, which was positioned in the neck region. The system was validated using videofluorography (VF), and it was concluded that high correlation existed between the results from the proposed mechanism and VF. However, the paper did not consider artifacts such as motion and speech.

2.2.3 Non-Invasive Diet Monitoring using Infrastructural Sensors

Saeki *et al* [47] proposed a measuring system of food intake using image processing method as shown in Figure 2-7. The system was composed of 4 incandescent lamps, a USB

camera with 320×240 resolution and a computer running the proposed algorithm. A tray with plates and bowls on was taken a picture by the system before and after the food intake experiments. The software running on the computer consisted of two parts, i.e., the image processing program, and the data base program (DBP). The image processing program included the communication program, photography processing program, and the measuring processing program, while the DBP was composed of the dish database, food menu database, food database, food stuff database and personal database. When comparing the images before and after the experiment, the nutrition intake was calculated by referring to the detailed information from the database. This method has the advantage of being able to estimate the energy intake directly and not requiring any wearing devices. However, it also has some drawbacks: (1) the system is not portable, so that it cannot be used when the subjects eat outside; (2) the accuracy is expected to be much lower if the food is layered, as the camera only captures the image from the top; (3) the usage of database is questionable, because there are too many kinds of food around the world to be included completely, and people may even cook at will.



Figure 2-7: Camera based system proposed in [47]

Chang *et al* [48] designed and implemented a diet-aware dining table that could track the type and volume of food the subjects had taken as shown in Figure 2-8. The dining table was

augmented with two layers of weighing and RFID sensor surfaces, where the RFID tag at the bottom of the container indicated the type of food, and the weighing sensor measured the changes in weight. A weight-RFID matching algorithm was proposed to detect and distinguish how people eat. Experiments were carried out to validate this method including scenarios such as live dining (afternoon tea and Chinese-style dinner), multiple dinning participants, and concurrent activities chosen randomly. An accuracy of 80% was reported through the experiments. This work is able to report the per-dinner energy intake, however, it has the following drawbacks: (1) it can only monitor energy intake when people always have food on the table; (2) food in the container is sometimes heterogeneous, for example, a dish may consist of low calorie vegetables and high calorie beef, thus measuring the weight of the whole dish may not indicate how much energy the subject has taken; (3) the system has the limitation that dishes should be placed in rather than cross cells where RFID antenna and weighing sensors were located, and subjects should not place their elbows or hands on the table.



Figure 2-8: The embedded RFID and weighing table surfaces

2.3 Proposed Mechanism

This dissertation presents the design, system level details and algorithms of a wearable food and drink intake monitoring system that analyzes human breathing signal. Food and drink intake can be detected by the way of detection a person's swallow events. The system works based on a key observation that a person's otherwise continuous breathing process is interrupted by a short apnea when she or he swallows as a part of the intake process. We detect the swallows via recognizing apneas extracted by a wearable sensor chest-belt. Such apnea detection is performed using matched filters and machine learning mechanisms, and further refined using a Hidden Markov Model (HMM) based mechanism that leverages known locality in the sequence of human swallows. This dissertation also demonstrates the effectiveness of the proposed mechanisms using experimental data.

Comparing to the existing work using non-invasive sensors for food intake monitoring, our work has the following advantages:

- (1) Ease of Usage: the work [28]- [37] attached sensors on the neck region for collecting swallowing sounds using elastic bands or adhesive tapes, however, the elastic bands and wiring will cause cosmetic issue so that people may be reluctant to wear those systems. While our system used piezoelectric belts or RIP belts for collecting breathing signal, which can be worn inside or between garments, and no sign of wearing can be observed. Moreover, since the piezoelectric belt works solely on a small piece of piezoelectric sensor, which can be even embedded into clothes, it is very suitable for long-term diet monitoring.
- (2) Usage Comfort: in our experiments [49], a microphone and an elastic belt were initially used as control, however, subjects complained that the belt was uncomfortable and may affect their swallowing patterns, i.e., they may have more spontaneous swallows. Therefore, the long-term usage of microphone on neck region fixed using elastic belts can be questionable. In some other work [29] [39] [40] [41], SEMG

electrodes were used to collect EMG signal on the skin in the neck region, however, in order to provide reliable contact, conductive gel was normally used, which may contaminate the clothes and not comfortable for everyday usage.

Chapter 3: Instrumentation System

This chapter first introduces the concept of apnea during swallow, which is the key observation that we leveraged for intake monitoring, and then compare a number of different sensors that can be used for breathing signal extraction in our system. Lastly, our proposed system is described in detail

3.1 Breathing Process

Breathing is the process during which air moves in and out of the lungs to deliver oxygen and remove carbon dioxide. The lungs can be expanded and contracted in two ways: (1) the human diaphragm moves up and down to lengthen or shorten the chest cavity, and (2) the ribs move back and forth to increase and decrease the anteroposterior diameter of the chest cavity [50]. It has been found that for healthy subjects, the movement of the rib cage contributes to around 75% of the tidal volume for both resting and exercising scenarios, while tidal volumes increase with the intensity of exercise [51]. For healthy people, chest and abdomen movements are synchronous with tidal air flow, but for some patients with chronic obstructive pulmonary disease (COPD), chest movement is synchronous with flow of air, but the abdomen moves asynchronously during parts of the breathing cycle [52].

3.2 Swallowing Apnea

Anatomically, breathing is inhibited in a part of the swallow process, thus causing a swallow apnea. The swallowing process is divided into three steps [53]: 1) the oral preparation phase, 2) the pharyngeal phase, and 3) the esophageal phase as shown in Figure 3-1. During the oral phase, food is chewed into a viscous bolus, and liquid is also considered as bolus with very high fluidity. The volume and viscosity of bolus is also sensed in this phase, so that the
swallowing apparatus can adapt to the bolus. In the pharyngeal phase, the bolus travels through the pharynx and passes the upper esophageal sphincter. A set of muscles are activated to propel the bolus and the epiglottis moves downward to cover the vocal folds and to protect the trachea from contamination. Finally, the bolus is pushed towards the stomach during the esophageal phase. During the pharyngeal phase, since the trachea is blocked by epiglottis, breathing is temporarily stopped, thus causing the apnea.



Figure 3-1: 3 steps of swallowing

Figure 3-2 shows the breathing signal of two subjects. The rising edges correspond to inhalations and the falling edges correspond to exhalations. As shown in the figure, a breathing cycle can be either normal (i.e. Normal Breathing Cycle or NBC) or elongated due to swallow-triggered apnea. A cycle that is elongated due to an apnea at the beginning of an exhale (see the top figure on the left in Figure 3-2 for subject-1, session-1) is termed as Breathing Cycle with Exhale Swallow (BC-ES). For a second subject, the bottom figure on the left in Figure 3-2 shows swallows (i.e. apnea) during the inhale process which are termed as Breathing Cycles with Inhale Swallow (BC-IS).

Swallow apnea localization with regard to breathing phases has been investigated in this dissertation. Martin *et al* [54] reported that swallow apnea happened during expiration with a probability between 0.94 and 1, and the swallow apnea was followed by expiration with a probability of 100% for 3, 10 and 20ml bolus size experimented. Klahn and Perlman [55]

reported similar results where swallow apnea occurs during expiration 93% of the time and was followed by expiration with probability 1 in their experiment with 5ml of water and applesauce. As a comparison, it was reported in [56] that the exhale-swallow-exhale pattern was 77% out of 100 swallows. Therefore, BC-ESs as shown in Figure 3-2 are much more common than BE-IS, which is validated in our following experiments.



Figure 3-2: Breathing signal of two subjects in the experiments

By detecting swallow events, the food or liquid amount can be estimated. Nilsson *et al* [57] has reported that the average bolus volume for 292 healthy adults during single swallows is 25.6 ± 8.5 ml and 21.1 ± 8.2 ml during repetitive swallows. When gender was considered, they demonstrated that the mean bolus size for single swallows is 28.1 ± 9.1 ml for males and 21.6 ± 5.5 ml for females, and for repetitive swallows, 23.2 ± 9.2 ml for males and 17.9 ± 4.8 ml for females.

3.3 Sensors for Collecting Respiratory Signal

Generally, there are some popular non-intrusive breathing detection techniques for obtaining a respiratory signal. Dietz [58] proposed a method using a flexible tube worn on the chest of the subject and connected to an external equipment measuring the airflow through the tube. Corbishley et al [59] used a miniaturized microphone in an aluminum conical bell on the neck region, and they proposed algorithms to handle noises from the body, the sensor itself and environment. Chia et al [60] developed a UWB radar based system, which is able to detect breathing and heart beat remotely even through walls. Mukai et al [61] used a 40kHz ultrasonic transmitter and receiver installed into a bed mattress for monitoring respiration, cardiac vibration and movement. Masks are also widely used for extracting breathing signal, and by detecting the oxygen concentration, they are able to estimate the energy consumption of human subjects[62] [63]. Karlen *et al* used ECG and photoplethysmography (PPG) to estimate respiratory rate. A pressure sensor array was used in [64][65] to detect breathing signal during sleep and a algorithm was proposed to select proper sensor sets during movement in [65]. Bates et al [66] Jin et al [67] Hung et al [68] and Reinvuo et al [69] used accelerometer to detection breathing signal and proposed algorithms to handle the artifacts caused by body movements and speech. Koo et al [70] proposed using a piezoelectric polyvinylidene fluoride (PVDF) impedance sensor for measuring respiratory signal and compared its performance with respiratory inductance plethysmography (RIP) and nasal-oral pneumotachography. Conductive rubber [71] can also be used to measure respiratory signal. The sensor is made of a string of carbon black and carbon filled silicone rubber, whose resistance changes when it is stretched.

Comparing to other sensing methods, conductive rubber, piezoelectric PVDF and respiratory inductance plethysmography (RIP) have the advantages of (1) non-intrusiveness:

these sensors can be put on and off easily, (2) no skin contact: these sensors can be put between garments without causing any cosmetic issue, and (3) comfort: these sensors can be embedded into garments and worn all day long without causing any discomfort.

In the following section, we compare the conductive rubber, piezoelectric PVDF and respiratory inductance plethysmography (RIP) solutions for respiratory signal detection, and discuss their advantages and disadvantages.

3.3.1 Conductive Rubber Sensor

An elastic belt on the chest or abdomen is able to capture the change in tension when the chest or abdomen expands or contracts. The tension is directly reflected as the change in the resistance, and it can be easily converted to voltage when put serially with another constant-value resistor.



Figure 3-3: Resistive stretch sensor (a) and its static response (b)

In order to evaluate characteristics of the resistive belt, we choose the stretch sensor manufactured by Scientific Instruments as shown in Figure 3-3:(a). The diameter of the sensor is 1.5mm, and the length is 15cm.

Static and transient property of the sensor is analyzed by stretching the sensor to different length. Figure 3-3:(b) demonstrates the static property of the sensor. In this experiment, the resistance of the sensor at each sample point is read 1 minute after it is stretched so that the impact of transient response is minimized. From Fig2, the sensor demonstrates good linearity in static experiments.

To analyze the transient property, the sensor is first stretched for 5cm to make it tight so that the impact of slack is minimized. Then the sensor is stretched for another 5 cm, and the response is therefore recorded. After 1 minute, when the resistance of the sensor is stable, the sensor is release for 5 cm. Figure 3-4 shows the resistive response of the sensor when such an experiment is done.



Figure 3-4: Transient response of stretch sensors

From Figure 3-4, when the sensor is stretched, a jump can be observed followed by decay. Similarly when the sensor is released, its resistance jumps up quickly before decaying to its resulting resistance.

Generally, the resistive belt provides a cheap solution for breathing detection. However, its transient response deforms the breathing signal despite its linearity in static case.

3.3.2 Piezoelectric PVDF Sensor

The device contains a piezoelectric sensor placed between two elastic strips. Stretching the belt exerts a strain on the sensor, which generates a voltage proportional to the strength of the force. Compared with other transduction principles, such as capacitive, inductive and piezoresistive sensors, piezoelectric sensors provide highest sensitivity [72] and excellent linearity over a wide amplitude range.



Figure 3-5: Piezoelectric belt (a) and its static response (b)

In order to evaluate the characteristics of the Piezo respiratory sensor, both static and transient response analysis is done as for resistive belt. Figure 3-5:b shows the static response of the piezo respiratory sensor, and the sensor exhibits good linearity in the experiment.

The piezo electric devices have very high input impedance, and the equivalent circuit is shown in Figure 3-6. As the input resistances of normal oscilloscopes are around 1M Ω , and the Cs of the piezo respiratory belt we are using is 2.2µF, when the output of the sensor is directly injected into the oscilloscope, the time constant of the circuit is 2.2 second. Therefore, in order to analyze the transient response of the sensor when it is dragged or released, we need an isolate circuit for a larger time constant. Operational amplifier MAX406 is able to provide ~1011 Ω

input impedance, which would bring up the time constant to 2.2×105 seconds. Consequently, and isolation circuit shown in Figure 3-6:b is adopted to separate the oscilloscope and the sensor.



Figure 3-6: Equivalent circuit of piezoelectric sensor (a), Isolation circuit (b), and Transient response of piezo respiratory sensor(c)

Figure 3-6:c demonstrates the signal captured by the oscilloscope when the piezo respiratory sensor is stretched for 15mm and then released. The output signal is very clean and follows the mechanical input closely. Compared with resistive belt, the piezo respiratory belt has much better characteristics.

3.3.3 Respiratory Inductance Plethysmography (RIP) Sensor



Figure 3-7: RIP sensor

An inductance belt is relying on Farady's Law and Lenz's Law that a magnetic field will be generated when current flows through a loop of wire, and a change in the area enclosed by the loop would create current in the loop in the opposite direction proportional to the change in the area. When the inductance belt is used, a low amplitude sine wave of ~20 mV at ~300 KHz is injected through the belt. The inhalation and exhalation would change the area enclosed by the belt introducing an opposing current in the belt and thus deforming the applied current and changing the frequency. The frequency is then demodulated to produce analogy waveform reflecting the change of the area. It is reported that the output of the belt changes linearly with the cross-sectional area [73]. A RIP respiratory system is normally composed of two RIP belts, wearing on the chest and abdomen respectively.

Consequently, a driver module is needed to measure the breathing signal, which consists of a frequency generator, signal processor and analog/digital converter. The price of an inductance belt set is therefore much higher than the other two sensors analyzed.

As a conclusion, the conductive rubber sensor is inferior in terms of transient response. The piezoelectric belt provides a better solution with good linearity and transient response, yet its position needs to be adjusted for each subject for optimum signal amplitude. RIP belt system provides the best signal quality and does not need position adjustment as it has a dedicated signal processing driver module and two belts wearing on the chest and abdomen individually, although it cost much higher than piezoelectric belts. Therefore, in this dissertation, we have done experiment using both piezoelectric belts and RIP belts for swallow detection. For the RIP belt, we used the sum output of the two belts, i.e., on the chest and abdomen.

3.4 Swallow Detection System

Based on the analysis from Section 3.2, we developed two systems using piezoelectric belt and RIP belts respectively. In the experiment carried out in Chapter 4, 5 and 6, piezoelectric sensor belt is used, while RIP belt is adopted in our future experiments.



Figure 3-8: System architecture of swallow detection system using piezoelectric belt

The system architecture of swallow detection system using piezoelectric belt is shown in Figure 3-8. The embedded wearable sensor system is worn on the chest for collecting breathing signal and transmitting it to a smart phone through Bluetooth. The embedded belt system contains: 1) a piezo-respiratory belt for converting the changes of tension during breathing to a voltage signal, 2) an amplifier and signal shaping circuit for formatting the raw voltage signal to a format suitable for the ADC chip, 3) a processor and radio subsystem (EZ430-RF256x from Texas Instrumnet), and 4) a 3.7V 300mAH polymer rechargeable battery. The entire package weighs approximately 40 grams. The 300mAh polymer battery is able to support the system for more than 30 hours of continuous operation on a single charge. After the signal is received by the smart phone, it is stored on an SD card attached to the phone.



Figure 3-9: System architecture of swallow detection system using RIP belts

The swallow detection system using RIP belts is shown in Figure 3-9. Similar to the system based on piezoelectric belt, it includes: 1) a pair of RIP belts for collecting breathing

signal, 2) a signal shaping circuit for amplifying and filtering the raw signal from the sensor to cater the requirement of ADC stage, 3) a processor and Bluetooth subsystem (TI EZ430-RF256x), which is able to sample the signal at 100Hz and transmit it over Bluetooth to the external smart phone, and 4) a 3.7V 300mAh polymer rechargeable battery. The whole packet weighs approximately 45 grams. The 300mAh polymer battery is able to support the system for around 20 hours.

Due to the fact that the output signals from piezoelectric belt and RIP belts are very different in terms of output impedance, signal amplitude and signal to noise ratio (SNR), signal shaping circuits need to be designed individually to cater their characteristics. However, battery, μ Controller and Bluetooth module, app on smart phones can be shared between the two systems.

3.4.1 Signal Shaping Circuit

The out of these sensors are noisy and weak, for example, the output of the piezoelectric belt on a subject during normal breathing is only about 10mV, while the corresponding output for RIP belt is only 20μ V. Moreover, since the RIP respiratory belts are using high frequency current to drive the belt as discussed in 3.3.3, the output signal is also contaminated with those high frequency components.

A. Signal Shaping Circuit for Piezoelectric Belts

Signal shaping circuit is needed for the piezoelectric belt for two reasons. First, based on our observation, when people are breathing, the change in the circumference of the chest is normally within several millimeters, which makes the output of the belt by fluctuating within 10 mV. In order to capture this signal using main stream ADCs on mobile sensors, amplification circuit is necessary. Second, the output signal of piezo respiratory belt is proportional to the tension on the belt, so when the belt is worn on the chest or abdomen, a certain amount of electrons will be accumulated on the two sides of the sensor because of the tension. The electrons will become the DC part of the respiratory signal collected on the belt. This DC value may change from person to person depending on how they wear it, and it may also change from time to time when the tension on the belt changes slowly. The uncertainty of the DC value makes the design of the amplification circuit complicated. Therefore, this amplification circuit must be designed carefully to meet these two requirements.



Figure 3-10: Signal shaping circuit for piezoelectric belt

Amplification circuit as shown in Figure 3-10 is designed to meet the two requirement mentioned above. The circuit can be divided into four parts. The first part is drifting control. It adds an $8.2M\Omega$ resister to the output of the belt, which is able to damp the DC drifting when the belt is worn by different people. The time constant introduced by the $8.2M\Omega$ resistor is 18.04

seconds, which would reduce the impact of DC drifting while bringing minimum effect on respiratory signal. The second part is voltage shifting. It sets the default DC voltage to be 63mV, so that the output signal stays positive. The third part is impedance match. It isolates the previous parts from the amplification circuit. The last part is amplification. It is consisted of a low pass filter and an amplification circuit. The time constant of the low pass filter is 5.6 milliseconds, and it can filter out the noise caused by previous circuit. The amplification rate of the circuit is controlled by an adjustable resistor.

B. Signal Shaping Circuit for RIP Belts

Since the output signal amplitude of the RIP sensing module is very small (around $20\mu V$ peak-to-peak for normal breathing) during normal breathing, high amplification ratio is necessary for the signal shaping circuit. A low pass filter is also implemented to filter out the high frequency components in the breathing signal caused by the alternating current injected into the belts. Figure 3-11 shows the schematic of the circuit. The first op-amp (i.e., MAX412 (1) in the figure) is to provide a stable 1.5V reference to the input and the amplifier on the right. The second op amp (i.e., MAX412 (2) in the figure) forms the amplifier with an amplification ratio of 10,000. The low pass filter accompanying the amplifier has a cutoff frequency of 28Hz. After the amplification, the breathing signal has a peak-to-peak value around 200mV.

The MAX412 (MAX412BCSA+) op-amp is chosen because of its low input offset voltage and low input noise-current density, i.e., the input offset voltage $V_{OS}=\pm 150\mu V$, and input noise-current density $i_n=1.5nV/Hz^{1/2}$.



Figure 3-11: Schematic of signal shaping circuit

3.4.2 µController and Bluetooth Module

The processed signal from the signal shaping circuit module is fed into an ADC channel of a μ Controller and Bluetooth module (EZ430-RF256x). The ADC converter has an accuracy of 14 bits, and a sampling rate at 100Hz. For a breathing signal with 200mV peak-to-peak value, the SNR regarding the quantization noise is 60dB.

The sampled data is then sent via Bluetooth to the Android smart phone for food intake monitoring application.

In addition to the breathing signal from the belts, acceleration data collected from the accelerometer on the μ Controller and Bluetooth module is also collected and transmitted to the smart phone for future usage.

3.4.3 Breathing Signal Logging App

An Android app, named LiveActive, is developed on the smart phone to perform food intake detection based on breathing signal received from Bluetooth.



Figure 3-12 demonstrates the logic layering of the LiveActive. The LiveActive activity coordinates the other modules. Bluetooth connection module receives breathing signal from Bluetooth API provided by Android. Graphic User Interface (GUI) plots the breathing signal on the screen. Food intake detection module runs in background to perform the detection algorithm illustrated in Section 3. Webservice module updates the detection results to the server using either WIFI or cellular network. Currently, the food intake detection module and webservice module are still under development, and the collected data are stored in the microSD card in the smart phone for further processing. The proposed design of the webservice module is shown in Figure 3-13.



Figure 3-13: Design of the webserve module



Figure 3-14: Graphic User Interface (GUI) of LiveActive

As shown in Figure 3-13, the breathing signal is first received through the Bluetooth interface as a constant data stream, which will then be packetized into data packets. The size of the packet is going to be optimized to achieve the shorter transmission delay and less memory usage. Json serializer is used to convert the data packet into a data array, which can be

transferred through the Ethernet to the server side. On the server side, the data is converted back to a continuous data stream, and the intake detection algorithm is launched to detect swallow events embedded in the breathing signal.

Figure 3-14 shows the GUI of the app LiveActive. The GUI is mainly divided into three parts. The first part sits on the top of the screen showing the raw data received from Bluetooth interface, the second part plots the breathing signal, and the last part plots the acceleration data of three axes. All the three parts get updated when a new sample point is received.

Chapter 4: Swallow Detecting using Matched Filters

The sensor system and detection algorithms developed in this chapter works based on a key observation that a person's otherwise continuous breathing process is interrupted by a brief apnea during a swallow, which is a part of the intake process [74]. We first detect swallows by the way of detecting apneas extracted from breathing signal captured by a wearable wireless chest-belt with piezoelectric sensors as demonstrated in Chapter 3.4. Afterwards, swallow pattern analysis is used for identifying drinking swallows. Together with self-reporting at the high level of overall liquid intake habits (i.e., the types of drinks etc.), the instrumented detection of liquid swallow counts can offer an objective way to: 1) study the liquid intake trends, and 2) estimate calorie intake.

4.1 Processing Methods

Matched filters [75] are widely used in signal processing. A matched filter is implemented by correlation a known signal, or reference, with an unknown signal to detect the presence of the template signal in the unknown signal. This method is equivalent to convolving the timereversed reference signal to the unknown signal. The matched filter is the optimal linear filter, which maximizes the signal to noise ratio (SNR).

In this dissertation, we used matched filters to detect the presence of swallow apnea for swallow detection purpose.

4.1.1 Breathing, Apnea and Swallow Signature

Figure 4-1: Demonstrates two representative human breathing signal segments. The ADC readings in the figure are directly proportional to the elongation of the piezo-electric sensing belt shown in Chapter 3. A breathing cycle can be either normal (i.e., Normal Breathing Cycle

or NBC) or elongated due to a swallow-triggered apnea. A cycle that is elongated due to an apnea at the beginning of an exhale (see Figure 4-1:a) is termed as Breathing Cycle with Exhale Swallow (BC-ES). Figure 4-1:b shows swallows (i.e., apnea) during the inhale process which are termed as Breathing Cycles with Inhale Swallow (BC-IS). During our experiments, it was also found that BC-ES is much more prevalent than BC-IS, which also coincides with previous research in [54][55][56].



Figure 4-1: Examples of Breathing Cycles (BC), Normal Breathing Cycles (NBC), Breathing Cycles with Inhale Swallow (BC-IS), Breathing Cycles with Exhale Swallow (BC-ES) and apnea

The objective in this chapter is to be able to classify NBC, BC-ES, and BC-IS with high accuracy. The challenges in detection stem from the fact that there is significant variability in breathing waveforms across different: 1) subjects, 2) measurement instances for the same subject, and most importantly, 3) the location and duration of the apnea with respect to its

breathing cycle. Moreover, breathing signal may also be contaminated by artifacts such as movement, speech, coughing and etc.

4.1.2 Matched Filter Method

In what follows we present the performance of a matched filter based template matching mechanism for swallow detection. The template signals for matched filters are chosen from NBCs, BC-ESs, and BC-ISs, so that a breathing cycle can be classified as one of those three by observing the similarity score produced by the respective filters.

As shown in Figure 4-2, the signal sampled by ADC at 30Hz is first fed into a low-pass filter for removing any quantization noise. The second step is to extract individual breathing cycles through a peak and valley detector. The next module is for normalizing the extracted cycles in both time and amplitude, so that both input waveform and the reference waveforms of the matched filters have the same amplitude and the number of sample points. Each breathing cycle is normalized to be between 0 and 100 (ADC output units) vertically, and interpolated to 128 sample points. Considering the average length of a breathing cycle of 3.77 seconds in our experiments, the normalized sampling rate after interpolation is approximately 34Hz. Note that the time-normalization provides a way to handle variable duration breathing cycles and variable duration apnea (i.e., caused by different amount of liquid intake in one swallow) by creating a uniform duration swallow signature. Such a uniform duration signature is then presented to the proposed matched filter and machine learning algorithms.



Figure 4-2: Detection process of matched filter based detection algorithm

The normalized breathing cycle waveforms are fed into three separate filters, each with a specific type of reference template waveform. The filters use reference waveforms corresponding to NBC, BC-ES, and BC-IS. The similarity score outputs are compared in order to classify a breathing cycle as one of the above three types of BCs.

Note that the bottom part of Figure 4-2 shows how the shaped signal is used for feature extraction and swallow classification of using machine learning based methods as presented in Section 4.1.3.

4.1.3 Machine Learning based Detection

A machine learning approach with time domain features is applied using all 128 sample

points in a normalized breathing cycle. The Toolkit Weka [76] was used for implementing three different classifiers, namely, Support Vector Machine (SVM), Decision Tree (J48), and Naïve Bayes. The classifier parameters are optimized to provide the best accuracy. For SVM, polynomial kernel function is used and the features are normalized. All the other parameters are set to default values. A 10-fold validation approach is used in which the collected breathing cycles are randomly divided into 10 subsets of equal size and a classifier is run for 10 times. In each run, one subset is used for testing while the others are used for training. 10-fold validation method is used to avoid over-fitting [77].

Breathing signal power at different frequencies, computed using FFT, is also used as features in machine learning. Since FFT is applied on normalized 128-point breathing cycles with a normalized sampling frequency at 34Hz, it produces spectral coefficients for frequencies up to 34Hz with a granularity of 0.27Hz. As each normalized breathing cycle is a real finite length series, the resulting 128-point power spectrum are symmetric on $f_s/2$ [78]. Therefore, the first 64 spectral power values are used as the features for driving the classifiers.

4.1.4 Artifacts Handling

Breathing signals may suffer from artifacts, and in this dissertation, we proposed methods for handling artifacts caused by moving and speech.

Anatomically, it is not possible to swallow while talking, but people can talk right before or after swallowing. Therefore, it is necessary to detect talking so that swallow detection from breathing signal can be paused whenever talking is detected. Figure 4-3 demonstrates an exemplary breathing signal with speech artifacts. Observe that exhalation parts have larger slopes and more undulation, which is caused by modulated air flow through vocal folds.

In this dissertation, we proposed using power spectral density analysis to identify breathing

47

cycle with speech.



Figure 4-3: Breathing signal variability before, during, and after talking.

When people are eating or drinking while sitting, slight upper body movement may be involved. Therefore, exaggerate upper body movement was also involved in the experiments to analyze its impact on breathing signal.

4.2 Results

Experiments using the piezoelectric chest belt system were carried out with seven subjects (five male and two female) without any known respiratory or swallowing disorders. The belt was worn immediately inferior to sternum, where best signal strength was derived across all subjects. Each subject performed three sessions, 10 minutes each. In the first 5 minutes, the subject was asked drink water from a flask with a swallow instruction given once in every 20 seconds. Then the subject conversed with the experimenter for 3 minutes, and in the last 2 minutes, the subject shook their upper body and drank every 20 seconds. Breathing signals from the first 5-minute phase were used for swallow detection, and that from the last 5 minutes were for artifacts handling. First phase of each session resulted in approximately 80 Normal Breathing Cycles (NBCs) and 20 breathing cycles with swallows (both BC-ESs and BC-ISs). Please note that

spontaneous swallows were also included. For the first phase of each session, approximately 100 breathing cycles were recorded in total. For each subject, a library containing cycles from three such sessions (i.e., around 300 cycles) was then formed.

4.2.1 Results for Matched Filter based Detection

Templates or reference waveforms for the matched filters are computed based on cycles from the library as constructed above. A template for NBC is created by sample-by-sample averaging of three randomly chosen NBC breathing cycles from the library. Such randomness adds to the desired variability while forming the template. Similar process is adopted for constructing the templates for BC-IS and BC-ES breathing cycle types. One set of NBC, BC-IS and BC-ES cycles is referred to as a template combination.

Figure 4-4 shows the performance of matched filter based swallow detection method while using a large number of template combinations as the reference signals (i.e., S(n)) for all three breathing cycle types. For each subject, 3500 different template combinations are first created from the library by choosing different random combinations of NBC, BC-IS and BC-ES cycles while forming the templates. Then, each waveform in the library is classified to be one of NBC, BC-IS or BC-ES using the matched filter-based detection. An ROC pair (True Positive Rate, False Positive Rate) is finally computed for each of the 3500 template combinations. Figure 4-4 shows the resulting ROC distributions.



Figure 4-4: ROC Distribution for all seven subjects with arbitrary templates. The cluster of high value columns in Figure 4-4 indicates that even with arbitrarily chosen

template combinations, majority of them offer high True Positive and low False Positive rates. The spread in the distribution indicates that there exist NBC, BC-IS, and BC-ES waveforms which, if chosen to generate templates, can indeed degrade the system performance.

We experimented with different number of cycles (i.e., three for the above results) for generating the matched filter templates. It was observed that with more cycles, the spread in ROC performance were relatively shorter.

4.2.2 Performance for Machine Learning based Method with Time Domain Features

Classification accuracies for machine learning based methods using time domain features are summarized in Table 4-1 on both per-subject basis and on combined data set from all seven subjects. In the subject specific case, SVM and J48 perform better in terms of both True and False Positive rates. In the combined case, J48 provides the best swallow detection performance.

Subject	Classifier	True Positive Rate ±std (%)	False Positive Rate ±std (%)
Individual subject	SVM	98.69±2.03	0.14±0.16
	J48	98.8±0.49	0.32±0.14
	Naïve Bayes	97.7±2.62	2.63±2.3
Combined data set	SVM	87.6	1.5
	J48	97.5	0.7
	Naïve Bayes	85.5	9.5

Table 4-1: Performance of classifiers using time domain features

4.2.3 Performance for Machine Learning based Method with Frequency Domain Features

Table 4-2 shows the frequency domain detection performance of 3 classifiers on seven subjects using a 10-fold validation. Similar to the time-domain results, SVM and J48 provides

better True and False Positive rates for subject-specific classification. When classification is done on combined data from all seven subjects, J48 outperforms the other two.

Subject	Classifier	True Positive Rate±std (%)	False Positive Rate±std (%)
Individual subject	SVM	99.29±1.25	0.09±0.15
	J48	98.8±0.51	0.37±0.17
	Naïve Bayes	95.96±3.03	2.96±1.26
Combined data set	SVM	88.8	2.2
	J48	96.6	0.8
	Naïve Bayes	82.1	4.8

Table 4-2: Performance of all three classifiers using frequency domain features. Performance on both individual dataset and the combined dataset are presented in this table.

Note that in Tables 4-1 and 4-2, the standard deviation comes from the variation of true and false positive rates across the subjects. Since for the combined data set scenario in both the tables, the data from all seven subjects are combined into a single data set, standard deviation does not apply to that scenario.

4.3 Discussion

4.3.1 Iterative Template Refinement

The performance of matched filter based detection method heavily depends on the selection of reference waveforms. As shown in Figure 4-4, improper selection of reference waveform can deteriorate the performance significantly. In this section we develop an iterative template refinement algorithm for incrementally improving the swallow classification performance. First, an NBC, a BC-IS, and a BC-ES waveform are chosen from the breathing cycle library. Second, all collected breathing cycles are classified using those three waveforms as the templates to the three matched filters. At this stage, each collected cycle is classified as NBC,

BC-IS, or BC-ES. Third, all cycles that are classified as NBC are sorted based on their similarity score obtained from the NBC matched filter in the second step. Now, the top 50% of those NBC cycles are sample-by-sample averaged to create the NBC template for the second iteration. The same process is also executed for BC-IS and BC-ES to form the templates for the second iteration. The second and third steps are iteratively repeated till the breathing cycles selected in the third step for generating templates stabilize.

Stabilization is defined as when the differences between the matched filter similarity scores across consecutive iterations reduce below a pre-defined threshold, which in turn, dictates the overall error performance of the mechanism. The algorithm is summarized in Algorithm 1.

Algorithm 1: Iterative template refinement algorithm

Input: Initial templates T_{NBC} , T_{BC-IS} , and T_{BC-ES} while (templates have not converged) for (all collected breathing cycle x_i) $\begin{bmatrix} for (all collected breathing cycle <math>x_i$) $T_{BC-ES} = average (detected BC-ESs with top 50\% \mu_i^{BC-ES})$ $T_{BC-IS} = average (detected BC-ISs with top 50\% \mu_i^{BC-ES})$ $T_{BC-IS} = average (detected BC-ISs with top 50\% \mu_i^{BC-IS})$ $T_{BC-IS} = average (detected BC-ISs with top 50\% \mu_i^{BC-IS})$ $T_{BC-IS} = average (detected BC-ISs with top 50\% \mu_i^{BC-IS})$ $T_{BC-IS} = average (detected BC-ISs with top 50\% \mu_i^{BC-IS})$ $T_{BC-IS} = average (detected BC-ISs with top 50\% \mu_i^{BC-IS})$ $T_{BC-IS} = average (detected BC-ISs with top 50\% \mu_i^{BC-IS})$

The key concept here is that even when the initial matched filter template quality is poor, by choosing the top 50% of the cycles, the algorithm is able to iteratively refine the template quality, thus delivering good final detection performance. Figure 4-5 depicts the algorithm dynamics in the form of the similarity score state space at the start and at stabilization of template refinement. The top graph shows the location of all the collected breathing cycles in the similarity score space obtained from the starting template waveforms. The graph has 247 points, corresponding to 247 collected breathing cycles. The bottom graph corresponds to similarity score space obtained from the template waveforms when the algorithm stabilizes.

Observe that the overlapping among the three types of breathing cycles is much less in the bottom graph compared to the top one. This indicates a clear improvement of the matched filter template quality, leading to improved separation of different classified cycle types. The tighter clustering of the points in the bottom graph provides additional indication to better template quality compared to the starting set. The patterns in Figure 4-5 have been consistently observed for a wide range of initial template quality applied to the data from all seven subjects.

Figure 4-6 shows the representative performance of iterative template refinement for a specific subject (i.e., subject-2). The evolution of true and false positive rates are reported for three different starting template sets, termed as, Good Starting Point (GSP), Moderate Starting Point (MSP), and Poor Starting Point (PSP). GSP represents the NBC, BC-IS, and BC-ES combination in the breathing cycle library that provides the highest true positive rate and the lowest false positive rate as evaluated in Figure 4-4. PSP, on the other hand, represents the combination in the library that provides the lowest true positive rate and the highest false positive rate. Finally, MSP is chosen to be a combination for which the true positive and false positive rates are somewhere in between.



Figure 4-5: Similarity score space for: a) initial matched filter template used as a starting point, and b) the final template obtained at stabilization of the iterative algorithm. The tighter clustering of the points in the bottom graph indicates iterative improvement of the template

quality.

Observe that the true positive rate for PSP consistently improves with iterations. For MSP, such rates either improve or remain constant. With GSP, true positive rates go down slightly,

although the decrement is always observed to be much less than the improvements observed for PSP, thus establishing the effectiveness of the approach.



Figure 4-6: Iterative template refinement performance; a) true positive rate, and b) false positive rate evolution with iterations.

Note that for few PSPs with highly deformed BC-ES or BC-IS waveforms, the false positive rates temporarily go up with iterations before they settle down to lower values. This explains the temporary increase in the false positive rate in Figure 4-6:b. For the majority of the PSPs, however, the false positive rate remains acceptably low. Results with waveforms from other subjects demonstrated very similar performance patterns.

4.3.2 Discrimination Power of Time Domain Features

The results of discrimination power analysis of the time-domain features (i.e., all 128 sample points of a breathing cycle) are shown in Figure 4-7. Figure 4-7:a depicts the overall importance of each feature in swallow classification in terms of merit. The merit here refers to information gain [79], which is defined as the reduction in class entropy (i.e., H(*)) with additional information provided by the feature about the target classes. Assuming A is the

feature, and C is the set of classes, the following two equations indicate the class entropy before and after providing the feature:





figure.

A feature with higher merit indicates lower class entropy when this feature is adopted. It also indicated higher utility of a feature, which can be used as a guidance factor when feature reduction is needed in the presence of limited computational and storage resources.

There are three distinct utility peaks in Figure 4-7:a, which can be explained using the breathing cycles shown in Figure 4-7:b. The sample points in peak region A are instrumental in distinguishing NBC from BC-IS and BC-ES, and those in region B help distinguishing BC-IS. Finally, the sample points in region C distinguish all three breathing cycle types. The implication of these results is that if a feature reduction is needed, unimportant samples can be eliminated

from the areas not in the vicinity of the peaks in Figure 4-7:a. While the results in Figure 4-7:a are for subject-1, we have observed very similar patterns of discrimination power for all seven subjects.

4.3.3 Discrimination Power of Frequency Domain Features

Figure 4-8:a depicts the overall importance of the spectral power at each frequency. The power in the frequency range 0 (i.e., DC) to approximately 3Hz contains the most information for differentiating the three target breathing cycle types. Figure 4-8:b reports the ROC graph with both time and frequency domain features, when only up to 5 features are allowed. The feature sets are selected using the subset evaluation method [79] as follows.



Figure 4-8: (a) utility of frequency domain features, (b) comparison between time and frequency domain features; results are presented with limited number of features that are chosen using a method as described.

The first feature is selected using the method illustrated in Section 4.3.2 which ensures the largest possible reduction in class entropy. The resulting first features are the 10th sample point (out of 128) in time domain, and the DC spectral power in frequency domain. These first features

in time and frequency domains can be observed in Figure 4-7:a and Figure 4-8:a respectively. Using the same procedure as above, the rest of the four features are added iteratively while maximizing the reduction in class entropy [79].

With more features, the difference in detection performance between the time and frequency domain approaches is negligible. When only one feature is used, however, the frequency domain approach outperforms the time domain approach for the following reason. Figure 4-8:a demonstrates that the DC component has the highest discriminative power, which can be expressed as:

$$X_{k=0} = \sum_{n=1}^{N-1} x_n \cdot e^{-i2\pi\frac{\kappa}{N}n} |_{k=0} = \sum_{n=1}^{N-1} x_n ,$$

where $X_{k=0}$ represents the area under curve of the breathing cycle waveform. For BC-ES, since the apnea is located at the beginning of an exhale, its area under curve is much higher than that of NBC and BC-IS. Moreover, majority of the swallows are found to be BC-ES, which is why $X_{k=0}$ alone can be used to achieve considerable detection accuracy. However, no single feature in time domain is able to provide similar discriminative power.

Observe that machine learning can provide higher detection accuracy compared to the matched filters, although they require a-priori training. The matched-filters, on the other hand, can achieve acceptable performance using the iterative template refinement algorithm as presented in Section 4.3.1.

It should be noted that both the presented mechanisms are subject-dependent and require personalized training. The subject-dependency stems from the wide diversity of the breathing signals across subjects and their inherently different breathing patterns. In spite of such diversity, however, the swallow signatures were found to be detectable through appropriate algorithm training as proposed for both the techniques.

4.3.4 Artifacts Handling

We analyzed the undulated exhalation of breathing cycles during talking using power spectral density (PSD). Figure 4-9 shows the comparison between PSD of breathing signals during talking (solid lines) and those during NBCs and swallows (dashed lines). The density is computed over 4.27-second windows to facilitate 128-point FFT for the 30Hz sampling rate (4.27 second=128pts/30Hz). Observe that the PSDs with talking contain many more variations between 0 and 2Hz mainly because of the undulations during exhalation as illustrated in Figure 4-3. This was consistently observed across a large number of subjects and sessions.



Figure 4-9: Power spectral density (PSD) of breathing signals with talking and without talking, when normal breathing or breathing with swallows are executed.

Using the variance of difference between the PSD of received breathing signal and the reference NBC signal, it is possible to identify talking so that swallow detection can be paused during talking. Detection of talking is accomplished by using a threshold (of variance of

difference), which can be either manually set or can be trained using variance of difference as a feature.

In order to analyze the impacts of upper body movements, in the last 2 minutes of each session, the subject shook their upper body and drank every 20 seconds to simulate changing postures and rocking, which often occurs during food and drink intake sessions. Figure 4-10 shows the breathing signals for a subject both with and without such upper torso movements while swallowing. Note that the signals belong to the same subject, but it is constantly observed for other subjects also.



Figure 4-10: Breathing signal: (a) with upper body rocking movement, and (b) without the movement By comparing the signals with and without artifacts, one can observe that such movement artifacts do not introduce any noticeable changes to the breathing signal, mainly due to the fact
that such movement does not change the circumvent of the chest area where the belt is placed. Because of this very minimal impact, the swallow signatures are well preserved, and can still be clearly discerned using the mechanisms described earlier in the chapter. Therefore, the proposed mechanism for swallow signature detection does work even when natural upper body movements are present during a swallow process.

4.4 Summary

This chapter reported the algorithm design for a wearable liquid intake monitoring system using piezoelectric chest belt illustrated in section 3.4. A matched filtered based template matching framework, along with a number of template design mechanisms, both static and iterative, were developed for swallow detection with high true positive rate and low false positive rate performance. This paper also presented the preliminary results for classifier based detection using both time and frequency domain features. Finally, talking and upper body motion artifacts were analyzed.

Please note that this chapter only focused on the detection of liquid intake, although swallowing apnea is also present for solid food intake. Using the features extracted from breathing cycles, we may be able to distinguish solid and liquid intake using the proposed chest belts.

62

Chapter 5: Machine Learning Based Processing Algorithms

In this chapter, we present a swallow detection algorithm based on machine learning methods. The system works based on the key observation that during swallowing, because the trachea is blocked, a person is not able to breathe, thus causing a temporary apnea. Using the wearable piezoelectric chest belt introduced in Chapter 3, we detect swallows by the way of detection apnea captured by the chest belt. After the swallow sequence is recorded, a swallow pattern analysis can potentially be used for identifying non-intake swallows, solid intake swallows, and drinking swallows.

Comparing algorithms proposed in Chapter 4, where only liquid intake monitoring was tested, in this chapter, we demonstrate the algorithm and software extension of the same concept for monitoring both solid and liquid intakes.

5.1 Processing Methods

5.1.1 Machine Learning Algorithms

Machine learning is a branch of artificial intelligence, and it provides methods and algorithms for building systems that can learn from the data. For example, a machine learning system can be trained on a large number of hand written digits to recognize them, and after training, the system can be used to classify new hand written digits.

A number of machine learning algorithms have been developed, such as Artificial Neural Network (ANN), Support Vector Machine (SVM), Naïve Bayes, and Decision Tree.

<u>Artificial Neural Network</u>: An ANN is composed of a network of inter connected neurons in different layers [80][81]. A typical ANN has three layers, the first layer has input neurons, which sends data via synapses to the second layer of neurons, and then via more synapses to the third layer of output neurons. More complex systems will have more layers of neurons or more number of neurons in each layer. The synapses have parameters called weight that manipulate the data in the calculation. Each neuron has a transfer function which determines its output based on the weighted inputs. The output of a neuron can be expressed as:

$$y = F(\sum_{i=0}^{m} w_i x_i + b)$$

Where:

- x_i is the ith input value,
- w_i is the weight associated with ith input value,
- b is the bias,
- F is the transfer function,
- y is the output value.

When the structure and transfer function of an ANN are predefined, by adjusting the weights and biases in an ANN, we can train the network to produce the output we want for specific inputs.

<u>Support Vector Machine</u>: SVM is a classifier that constructs a hyperplane in a high dimensional space, which is able to classify the data with the largest separation, or margin between the two classes. For data sets that are not linearly separable, SVM maps the original space into a much higher dimensional space using a kernel function. The hyperplane is known as the maximum-margin hyperplane, and the classifier is also called the maximum margin classifier. The decision function can be expressed as:

$$D(\vec{d}) = \sum_{k=1}^{T} c_k a_k K(\vec{d}_k, \vec{d}) + b$$

Where:

 $(\vec{d}_k, c_k), \ k = 1, 2, ... T$ is the training set, \vec{d}_k is the features of instances in the training set, and c_k is the class label for training instance with feature \vec{d}_k

 a_k and b are parameters trained using the training data set.

In the case of classifying more than two classes, the method one-versus-all is normally used, in which one class is distinguished against all the other classes.

<u>Naïve Bayes</u>: Naïve Bayes classifier is a probabilistic classifier based on Bayes' theorem with the assumption of strong independence among the input features. The Bayes theorem shows:

$$P(c_j \mid \vec{d}) = \frac{P(d \mid c_j)P(c_j)}{P(\vec{d})} \quad \vec{d} \in \mathbb{R}^n, c_j \in \mathbb{Z}$$

Where:

 $P(c_j | \vec{d})$ is the probability of an instance with feature vector \vec{d} being in class c_j , and it is also known as *posterior*,

 $P(\vec{d} | c_j)$ is the probability of generating an instance with feature vector \vec{d} given class c_j , and it is also known as *likelihood*,

 $P(c_j)$ is the probability of occurrence of class c_j , and it is also known as *prior*,

 $P(\vec{d})$ is the probability of occurrence of an instance with feature vector \vec{d} , and it is also known as *evidence*,

In practice, as $P(\vec{d})$ does not depend on class c_j , in classification problem, it is a constant among classes, and it is therefore ignored. $P(\vec{d} | c_j)$ and $P(c_j)$ can be estimated based on the training data set. Therefore,

class of
$$\vec{d} = \operatorname*{arg\,max}_{c_j} P(c_j \mid \vec{d})$$

<u>Decision Tree</u>: A decision tree is a classifier expressed as a recursive partition of the feature space. A decision tree has 3 types of node:

- Root node: the root node is the entry point of a decision tree, which have no input and multiple output,
- (2) Test node: a test node has exactly one input edge, and runs *if-ifelse-..-else* statement, it may have two or more output edges,
- (3) Leaf node: a leaf node has one input edge, and it demonstrates the class that an instance belongs to.

We use J48 (also known as C4.5) in this dissertation. In each test node, J48 chooses the feature that most effectively splits the data set. The splitting criterion is the normalized information gain, which is discussed in Chapter 6. The feature with the highest normalized information gain is chosen to make the decision. The algorithm is then recursively applied on the smaller subsets.

Decision trees can get undesirably complex and over-fitting, such that each training instance takes one node. Therefore, the stopping criteria are used by decision tree algorithms. The typical stopping criteria include [82]:

- Number of cases in the node is less than a threshold
- Percentage of instance belongs to a class exceed a predefined limit in a node

• Depth of the tree is within certain limit



5.1.2 Breathing Apnea and Swallowing Signature

Figure 5-1: Examples of Breathing Cycles (BC), Normal Breathing Cycles (NBC), Breathing Cycles with Inhale Swallow (BC-IS), Breathing Cycles with Exhale Swallow (BC-ES) and apnea

Figure 5-1 demonstrates two representative human breathing signal segments. The ADC readings in the figure are directly proportional to the elongation of the piezo-electric sensing belt shown in Chapter 3. A breathing cycle can be either normal (i.e., Normal Breathing Cycle or NBC) or elongated due to a swallow-triggered apnea. A cycle that is elongated due to an apnea at the beginning of an exhale (see Figure 5-1:a) is termed as Breathing Cycle with Exhale Swallow (BC-ES). Figure 5-1:b shows swallows (i.e., apnea) during the inhale process which are termed as Breathing Cycles with Inhale Swallow (BC-IS). During our experiments, it was also found

that BC-ES is much more prevalent than BC-IS, which also coincides with previous research in [54][55][56].



Figure 5-2: Example breathing signals for solid and liquid swallows

Figure 5-2 shows example breathing signals with solid and liquid swallows. As can be seen, for solid swallows, breathings are deeper and contain more temporal fluctuations. The key objective is to be able to classify three types of breathing cycles, namely, NBC, BC-ES, and BC-IS, and to detect if the swallow is a solid or liquid one. The challenges stem from the fact that there is significant variability in breathing waveforms across different: 1) subjects, 2) measurement instances for the same subject, and most importantly, 3) the location and duration

of the apnea with respect to its breathing cycle. Among other things, this depends a great deal on the swallowing habits and the texture of the material that is being swallowed.

5.1.3 Detection Scheme

Figure 5-3 depicts the logic for classifying breathing cycles towards swallow detection. The raw data sampled by ADC at 100Hz is first fed into a low-pass filter for removing quantization noise caused by the A-to-D conversion process. Because the power spectrum of breathing signal is mainly below 2.5Hz, 100Hz is obviously sufficient. The second step is to run the filtered data stream through a peak and valley detection module in order to extract the individual breathing cycles. The next module is used for normalizing the extracted cycles in both time and amplitude dimensions. Each breathing cycle is normalized to be between 0 and 100 vertically, and interpolated to 128 sample points. Considering the average length of a breathing cycle of 3.77 seconds in our experiments, the normalized sampling rate after interpolation is mapped to 34Hz. The objective of normalization is to make sure that although different cycles may have different time and amplitude ranges (person-to-person or cycle-to-cycle for the same person), they can be effectively identified based on the apnea caused by swallowing.

The normalized breathing cycle waveforms are fed into a feature extraction module which extracts time domain or frequency domain features. These extracted features are then selected based on their discriminative power, and fed into a classifier for training or testing purposes. Number of features would affect the complexity and performance of classification. A classifier would be simple but with inferior performance if very few features are selected. Classifiers with a large number of features, however, are complex but do not necessarily provide superior performance [77].



Figure 5-3: Logic for swallow signature detection

A hierarchical classification scheme is used for solid and liquid swallow detection. The first classifier detects if a breathing cycle is an NBC or a breathing cycle with swallow. The second classifier detects if a swallow is a solid and liquid when the output of the first classifier is a swallow.

5.2 Experiments

Experiments using the piezoelectric breathing belt were carried out for swallow detection with three subjects, including 2 male and 1 female. Each subject performed three liquid swallow sessions and three solid swallow sessions, each session lasting for five minutes. Each subject was asked to wear the instrumented chest-belt and sit still throughout the experiment. During the liquid swallow session, the subject drank water from a flask with a swallow instruction given once in every 20 seconds. 20 ml of water was added to the flask for each swallow, ensuring the swallow volume to be 20ml. Each liquid swallow session resulted in approximately 80 Normal Breathing Cycle (NBC) and approximately 15 breathing cycles with swallows (both Breathing Cycle with Exhale Swallow (BC-ES) and Breathing Cycle with Inhale Swallow (BC-IS)). During the solid swallow sessions, the subject was asked to eat 6 grams of crackers each time at their comfortable rate, and noted the time when he or she swallowed. Considering that the cracker would be chewed and mixed with saliva, the formed bolus was roughly the same volume as 20 ml of water swallows. The resulting swallow signals are collected over the Bluetooth channel on a smart phone.

		NBC (Seconds)	Solid swallow (Seconds)	Liquid swallow (Seconds)
Subject 1	Maximum	5.61	6.81	7.56
	Minimum	2.36	2.91	3.36
	Average	3.24	4.86	4.79
Subject 2	Maximum	5.88	9	5.56
	Minimum	1.64	3.54	3.57
	Average	3.44	6.27	4.27
Subject 3	Maximum	4.51	9.33	6.64
	Minimum	1.93	4.26	2
	Average	3.05	6.22	4.27

Table 5-1: Durations of different breathing cycle types

Table 5-1 summarizes the duration of different types of breathing cycles. In addition to the spread of the cycle durations across subjects, it should be observed that the cycles with swallows

(i.e., both solid and liquid) are consistently longer than the normal breathing cycles. This is mainly due to the short apnea introduced by the swallow events. Moreover, it can also be observed that there is significant difference in the lengths of solid swallow and liquid swallow, which is mainly because of the different texture of the bolus in solid swallow and liquid swallow.

5.3 **Results and Discussion**

5.3.1 Feature Extraction and Selection

As analyzed in our previous work [83], both time domain and frequency domain features can perform well in detecting liquid swallows. The discriminative power of those feature types, however, can be different. As shown in Figure 5-4:a, for time domain features, sample points with indices near 16 and 90 are more important than other sample points in classification. As shown in Figure 5-4:b, for frequency domain features, lower frequency components have more discriminative power. It was also found that the discriminative power distribution of frequency domain features are more consistent across subjects, which is why the time-domain features are used in this dissertation.



Figure 5-4: Discriminative property of time and frequency domain features

The second set of classification features is derived from the first derivative of the breathing signal. As shown in Section III, it was found that the solid swallows generally create more fluctuations in the breathing signal compared to the liquid swallows. To capture such fluctuations, an additional classification feature was derived from the first derivatives of the breathing signal. More specifically, the number of ± 10 crossings is used as the feature, which is defined as the number of points in the breathing signal at which the first derivative of the signal is exactly ± 10 or ± 10 . Compared to the number of zero crossings, the number of ± 10 crossings not only captures the fluctuations observed in solid swallows, but also helps detecting the swallows in the first place. Figure 5-5 shows an example of the benefits of ± 10 crossings of first derivative is 1, which is the same as NBCs and is not sufficient in detecting the swallow, but the number of ± 10 crossings of the first derivative is 2 instead of 1 in case of NBC, which helps to detect the swallow.



Figure 5-5: Benefits of ±1 crossings as a classification feature

The third set of features is derived from various length distributions of the breathing cycles. Table 2 summarizes all the used features used in this paper.

	Features		
Frequency domain features	1 st order Fourier transform coefficient 2 nd order Fourier transform coefficient 3 rd order Fourier transform coefficient 4 th order Fourier transform coefficient 5 th order Fourier transform coefficient		
Features from waveform	Number of +10 crossings in first derivative Number of -10 crossings in first derivative		
	Breathing cycle length Inhalation length Exhalation length Inhalation depth Exhalation depth		

Table 5-2: Features selected for classification

5.3.2 Swallow Detection

All the above features are fed into the hierarchical classifier for solid and liquid swallow detection. In order to prove the generalizability, we adopt the leave-one-out method, in which case, data from all subjects are used for training except the one whose data is used for testing.

Table 5-3 and Table 5-4 report the performance of the hierarchical classifier using the leave-one-out method. As can be seen, SVM provides the best performance among all the applied methods for both the classifier stages. For the first stage, for all subjects the true positive rates remained higher than 82.9% and false positive rates lower than 1.6%. The performance of the second stage classifier has accuracy ranging from 88% to 73.33% when SVM is applied. Testing the system with more subjects is under way.

		True positive rate (%)	False positive rate (%)
Subject 1	SVM	82.9	1.6
	J48	76	2.4
	Naïve Bayes	100	1.2
Subject 2	SVM	84	0
	J48	88.6	4.9
	Naïve Bayes	97.1	4.1
Subject 3	SVM	86.7	0
	J48	83.3	8.6
	Naïve Bayes	93.3	8.6

Table 5-3: Performance of the first stage of the hierarchical classifier

		Accuracy (%)	
Subject 1	SVM	82.86	
	J48	80	
	Naïve Bayes	76	
Subject 2	SVM	88	
	J48	80	
	Naïve Bayes	68.6	
Subject 3	SVM	73.33	
	J48	70	
	Naïve Bayes	70	



5.4 Conclusion

This chapter reported the algorithm and performance of the machine learning based food and drink intake detection system. It presented the machine learning based swallow detection method using hierarchical classification scheme.

During the experiment and analysis it was found that food intake swallows have very regular temporal patterns in lunch session. Such temporal information can therefore be used for improving the system detection accuracy, which is going to be analyzed in the next chapter.

Chapter 6: Support Vector Machine and Hidden Markov Model based Processing Algorithms

This chapter presents a wearable solid food intake monitoring system that analyzes human breathing signal and swallow sequence locality for solid food intake monitoring. Food intake is identified by the way of detecting a person's swallow events. A Support Vector Machine (SVM) is first used for detecting such apneas in breathing signals collected from a wearable chest-belt. The resulting swallow detection is then refined using a Hidden Markov Model (HMM) based mechanism that leverages known locality in the sequence of human swallows. The chapter experimentally demonstrates the effectiveness of such two-stage SVM-HMM based mechanism for solid food intake detection via analyzing breathing signal and human swallow sequence locality.

In our previous work [83][84], we reported the effectiveness of such as system for detecting liquid-only intake monitoring using swallow signal analysis. As an extension to our previous work, in this chapter we focus on solid-only intake detection using a two stage SVM-HMM processing strategy as follows. After the swallow sequence is recorded, a Support Vector Machine (SVM) is first used for detecting such apneas in breathing signals collected from a wearable chest-belt. The resulting swallow detection is then refined using a Hidden Markov Model (HMM) based mechanism that leverages known locality in the sequence of human swallows. In a future publication we plan to report processing mechanisms and their effectiveness for joint liquid-solid intake monitoring.

The contributions of this chapter are: 1) combining SVM and HMM methods for processing breathing signals for solid food intake detection, and 2) experimentally demonstrating

77

the detection accuracy and effectiveness of the proposed system and the signal processing methods.

6.1 Processing Methods



Figure 6-1: Respiratory signal with swallow signature

The piezoelectric belt based breathing signal collection system proposed in Chapter 3 is used in this chapter. Figure 6-1 demonstrates a number of experimentally obtained breathing signal segments from different human subjects. The ADC readings in the figure are directly proportional to the elongation and contraction of the piezo-electric sensing belt. The rising edges correspond to inhalations and the falling edges correspond to exhalations. As shown in the figure, a breathing cycle can be either normal (i.e. Normal Breathing Cycle or NBC) or elongated due to swallow-triggered apnea. A cycle that is elongated due to an apnea at the beginning of an exhale (see the top figure in Figure 6-1 for subject-1, session-1) is termed as Breathing Cycle with Exhale Swallow (BC-ES). For a second subject, the bottom figure in Figure 6-1 shows swallows (i.e. apnea) during the inhale process which are termed as Breathing Cycles with Inhale Swallow (BC-IS).



Figure 6-2: Processing scheme for swallow detection

Figure 6-2 depicts the logic for classifying breathing cycles towards swallow detection. Before sending the data to the ADC, an anti-aliasing analog low pass filter circuit with cutoff frequency of 30 Hz is applied. The signal is then sampled by ADC at 100Hz and fed into a software-based low-pass filter for removing quantization noise caused during the A-D conversion. Because the power spectrum of breathing signal is mainly below 2.5Hz, 100Hz provides a fast enough sampling rate. The next step is to run the filtered data stream through a peak and valley detection software module in order to extract the individual breathing cycles. In order to perform peak and valley detection, the data stream is first divided into 30% overlapping 10-second windows, and then a threshold based algorithm from [85] is used. The threshold is set to $0.3(\max_{d(m)\in C} d(m) - \min_{d(n)\in C} d(n))$, where C is the set includes all the data points in the 10-second window, and d(m) and d(n) are the mth and nth sample points in the 10-second window.

After individual breathing cycles are extracted, they are normalized in both time and amplitude dimensions. Each cycle is normalized to be between 0 and 100 vertically, and interpolated to 128 sample points in time. Considering the average length of a breathing cycle of 3.77 seconds in our experiments, the normalized sampling rate after interpolation is mapped to 34Hz. Although different cycles may originally have different time and amplitude ranges (person-to-person or cycle-to-cycle for the same person), the normalization process removes such variance in duration and amplitude, thus making the cycles more suitable for the apnea detection process.

Feature extraction module takes breathing cycles before and after normalization and extracts features including: 1) breathing cycle length, 2) inhalation duration and depth, 3) exhalation duration and depth of breathing cycles before normalization, and 4) ± 10 crossing counts, 5) first 5 Fast Fourier Transform (FFT) coefficients of normalized breathing cycles. The details about the features extracted are demonstrated in Section 6.2.3. The features are then fed into the Support Vector Machine (SVM) detection module with posterior probability outputs, which are illustrated in more details in Section 6.1.2. At this stage, a posterior probability

indicates the SVM-detected probability of a given breathing cycle to be of types normal breathing or breathing with swallow. Information about swallow sequence locality is not utilized at this stage. Finally, the Hidden Markov Model (HMM) is applied to the posterior probability outputs of the SVM module to improve the detection performance by leveraging a-priori knowledge about swallow sequence locality. The HMM modeling is presented in Section 6.1.3.

6.1.1 **Two-tier Swallow Detection**

In our previous work [83][84], Support Vector Machine (SVM) was shown to be the best classifier for liquid swallow detection. Like in traditional usage of SVM [77], the classification output for each breathing cycle was a class label, which is normal breathing or breathing with swallow. After analyzing the classification errors in [83][84], it was realized that many of those errors can be corrected by applying known locality information in human swallow sequences. For example, people rarely swallow in many consecutive breathing cycles. Thus, whenever the classification output shows many consecutive breathing cycles, errors can be suspected and the misclassified instances can be identified/removed by applying higher level techniques such as the Hidden Markov Model (HMM). This motivates the two-tier detection using SVM and HMM presented in the next subsections.

6.1.2 SVM-based Swallow Detection with Posterior Probability

Consider the following training set of size T:

 $(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_T, y_T)$

In each training instance (x_i, y_i) , $x_i \in \mathbb{R}^n$ represents a set if n input features, and y_i is the corresponding class label. For a binary class system in our case, y_i can be defined as

$$\begin{cases} y_i = 1 & \text{if } x_i \in Breathing \ cycle \ with \ swallows \\ y_i = -1 & \text{if } x_i \in Normal \ breathing \\ i = 1, 2, ..., T \end{cases}$$

A traditional SVM decision function can be derived as [86]:

$$D(x) = \sum_{k=1}^{T} y_k a_k K(x_k, x) + b$$
(1)

where a_k and b are trained using the training dataset, T is the number of training instances, and $K(x_k, x)$ is the kernel function of SVM. Classification for a test feature set x_j using the decision function can be as follows:

$$\begin{cases} x_j \in Breathing cycle with swallows, & if D(x_j) > 0 \\ x_j \in Normal breathing, & otherwise \end{cases}$$

The distance between x_j and the decision boundary (i.e., that separates breathing cycle with swallows and normal breathing) with the maximum margin can be expressed as $D(x_j)/C$ [86], where C is a positive constant depending on a_k (k = 1, 2, ..., T), training feature set x_k (k = 1, 2, ..., T) and the kernel function. Therefore, $|D(x_j)|$ is positively correlated to the confidence of correct detection, meaning the closer to the decision boundary, the less confidence in correct classification.

In order for the HMM to be able to process the SVM output, the latter needs to be in the form of posterior probability ADDIN ZOTERO_ITEM CSL_CITATION {"citationID":"ptU5sHDL","properties":{"formattedCitation":"[32]","plainCitation":"[32]"},"cit ationItems":[{"id":591,"uris":["http://zotero.org/users/642418/items/TWVH4M6A"],"uri":["http://zotero.org/users/642418/items/TWVH4M6A"],"itemData":{"id":591,"type":"paper-

conference","title":"Food Intake Activity Detection Using a Wearable Microphone System", "page": "298 -301", "source": "IEEE Xplore", "event": "2011 7th International Conference on Intelligent Environments (IE)", "abstract": "A method for non-invasive monitoring of human food intake behavior and long-term dietary protocol has been developed by the sole use of chewing and swallowing sound sensors. A novel sensor system has been built containing an inear microphone and a reference microphone integrated in a hearing aid case in order to record chewing and swallowing sounds in the ear canal and environmental noise, respectively. Using manual labeled records of the food intake sounds of 40 participants we developed an algorithm to detect food intake activity in sound data. Comparison between sounds from both microphones enables discrimination the internal and between external sounds.","DOI":"10.1109/IE.2011.9","author":[{"family":"Passler","given":"S."},{"family":"Fis cher", "given": "W.-J."], "issued": {"date-

parts":[["2011",7]]}}],"schema":"https://github.com/citation-style-

language/schema/raw/master/csl-citation.json"} [32][86] as opposed to class labels used by traditional SVM models [77] as described above. An appropriately designed SVM [86] can indicate the probability that a given input feature set correspond to a specific class. This probability is referred to as the Posterior Probability for that class. In what follows we describe the mechanisms for computing such probabilities which are the input for the swallow sequence based Hidden Markov Model presented in Section 6.1.4.

Posterior probability for class-i is formally defined as $prob(class_i | input \ features) = prob(\pm 1 | x_i)$. This indicates the probability that a given input feature set x_i correspond to a breathing cycle with swallow or a normal breathing cycle. It follows that $prob(1|x_i) + prob(-1|x_i) = 1$. We use the following method for computing posterior probability using the SVM decision function $D(x_i)$, as proposed by Wahba in [87]:

$$prob(class_i \mid input features) = prob(y=1 \mid x) = \frac{1}{1 + \exp(A * D(x) + B)}$$
(2)

where A and B are constants and estimated by minimizing the negative log likelihood of training data set using regression methods.

6.1.3 Hidden Markov Model with Swallow Sequence Locality

The key concept of HMM in swallow detection is as follows. A sequence of breathing cycles are represented by a discrete time Markov Chain consisting of two states (i.e., normal breathing cycles and breathing cycles with swallows) that are hidden from an observer, meaning that an observer cannot directly determine which state the system is in at any given point of time. However, the posterior probability out of the SVM, which indicates the likelihood of the system being in any state, is visible to the observer. The idea of HMM formulation is that if the locality in swallow sequence dynamics and the mapping between the system's state and posterior probability observation are known (or measurable) to the HMM model, then by observing the posterior probability out of the SVM the current state in the Markov Chain can be estimated.

Hidden State Space: As shown in Figure 6-3:a, a breathing cycle sequence can be modeled as a hidden state machine with two hidden states, namely, Normal Breathing and Breathing Cycle with Swallows. The states are hidden because they are not deterministically known from posterior probabilities computed out of the SVM processing.



Figure 6-3: (a) Hidden breathing state machine and (b) HMM processing components

<u>Transition Probability Matrix</u>: It is defined as $A = \{a_{ij}\}$, where a_{ij} represents the probability of transitioning from state S_i to state S_j .

$$a_{ij} = prob(q_k = S_j \mid q_{k-1} = S_i)$$

It is assumed that q_k depends only on q_{k-1} , which means

$$prob(q_k | q_{k-1}) = prob(q_k | q_{k-1}, q_{k-2}, ..., q_1)$$

A is an 2×2 matrix for two breathing cycle types in our case. The transition probability matrix is constructed from the true swallow sequence detected by a video camera and push button. The probabilities in this matrix represent the swallow sequence locality information which is leveraged by the HMM processing

<u>Observation Matrix</u>: Although the states are considered hidden, the SVM-computed posterior probability at each state can be considered as an observable parameter for HMM modeling purposes. For a given state-i, the posterior probability $prob(y_i = 1 | x_i)$, generated by the SVM detector, is utilized for constructing an observation bitmap O_i in the following manner.

The probability range [0, 1] is divided into N equal windows (we used N=10) and each window is presented as a bit in the N-bit long bitmap O_i The bit corresponding to the window in which the posterior probability $prob(y_i = 1 | x_i)$ falls on is set to 1, and all other bits in O_i are set to 0. For example, with N=10 and $prob(y_i = 1 | x_i) = 0.71$, the observation bitmap for state-i will be $O_i = \{0,0,0,0,0,0,0,1,0,0,\}$

Now let b_{jm} be the probability that if an observation bitmap's mth bit is 1 (i.e., all other bits are 0s) then the system is in hidden state j. Formally stated:

$$b_{im} = prob(O = \{bit_1 = 0, ..., bit_m = 1, ...\} | State = S_i)$$

An observation matrix of size M×N (M: Number of states, N: number of bits in the observation bitmap) is constructed as $B = \{b_{jm}\}$. In this case, a 2×10 matrix is constructed by combing the true swallow events detected by a video camera and push button, and the SVM outputs $prob(y_i = 1 | x_i)$ after processing the chest belt sensor data. This observation matrix, together with the transition probabilities and the following initial probability array, is used for HMM processing as described in Section 6.1.4.

<u>Initial Probability Array</u>: The initial probability array is represented by a vector $\pi = [\pi_i]$ of length M (i.e. 2), in which

$$\pi_i = prob(q_0 = S_i) \quad 1 \le i \le M$$

 π_i indicates the probability that the initial state of the hidden state machine is S_i . Therefore, by definition

$$\sum_{i=1}^{M} \pi_i = 1$$

This array is formed using true swallow data gathered by the experimental system as described in Chapter 3. The swallow system as modeled by HMM can be expressed as a tuple: $\lambda = (A, B, \pi)$, where A, B, and π represent the hidden state transition matrix (i.e. known swallow sequence locality), the observation locality, and the initial condition of the state machine respectively.

6.1.4 HMM Processing

As shown in Figure 6-3:b, the processing model is fed by the HMM model $\lambda = (A, B, \pi)$ and the posterior probability observation sequence, and its outcome is an estimation of the current system state \hat{q}_t .

The probability of observing a sequence for a given model $prob(O | \lambda)$ can be expressed as [88]:

$$prob(O \mid \lambda) = \sum_{Q} prob(O \mid Q, \lambda) prob(Q \mid \lambda)$$
(3)

where $O = \{O_1 O_2 O_3 ... O_T\}$ is a sequence of observations and $Q = \{q_1 q_2 q_3 ... q_T\}$ is a sequence of states, and they have the same sequence length T. Note that q_i is the i-th state in the state sequence Q and it can represent any state $S_j, 1 \le j \le M$ (i.e. M=2). $prob(O | Q, \lambda)$ means the probability of having the observation sequence O given the state sequence Q and the model λ . $prob(Q | \lambda)$ indicates the probability of having the state sequence Q for the model λ . $prob(O | Q, \lambda)$ and $prob(Q | \lambda)$ can be expressed as:

$$prob(O | Q, \lambda) = \prod_{i=1}^{T} prob(O_i | q_i, \lambda)$$

$$(4)$$

$$prob(Q | \lambda) = prob(q_1) prob(q_2 | q_1) \dots prob(q_T | q_{T-1})$$

$$(5)$$

where $prob(O_i | q_i, \lambda)$ is the probability of observing O_i in state q_i , which corresponds to an element in the observation matrix B, $prob(q_1)$ is the initial probability of state q_1 corresponding to an relevant element in initial probability array π , and $prob(q_{i+1} | q_i)$ is the probability of transitioning from state q_i to state q_{i+1} , which corresponds to one element in transition probability matrix A.

By substituting $prob(O | Q, \lambda)$ and $prob(Q | \lambda)$ in Equation (3) using Equation (4) and (5),

$$prob(O \mid \lambda) = \sum_{q_1, q_2, \dots, q_T} prob(q_1) prob(O_1 \mid q_1, \lambda) prob(q_2 \mid q_1)$$
$$\dots prob(q_T \mid q_{T-1}) prob(O_T \mid q_T, \lambda)$$
(6) It can

be interpreted as follows. Initially the system is in state q_1 with probability $prob(q_1)$ and generates the observation O_1 with probability $prob(O_1 | q_1, \lambda)$. At the next time slot, the state transitions from q_1 to q_2 with probability $prob(q_2 | q_1)$ and produces an observation O_2 with probability $prob(O_2 | q_2, \lambda)$ Finally $prob(O | \lambda)$ is derived by summing the products over all possible state sequences.

The Forward-Backward Procedure [89] is adopted to simplify the calculation. Consider the forward variable $\alpha_i(i)$ defined as:

$$\alpha_{t}(i) = prob(O_{1}O_{2}...O_{t}, q_{t} = S_{i} \mid \lambda)$$

where $\alpha_i(i)$ indicates the probability of partially observing the sequence $O_1 O_2 ... O_i$ and state at time slot t is S_i given the model λ . It can be proved [88] through induction that:

$$prob(O \mid \lambda) = \sum_{i=1}^{T} \alpha_{T}(i)$$

Consider backward variable $\beta_t(i)$, defined as:

$$\beta_t(i) = prob(O_{t+1}O_{t+2}...O_T \mid q_t = S_i, \lambda)$$

which represents the probability that partially observing the sequence from time slot t+1 to

the end given the current state S_i and the model λ . Now another variable is defined as:

$$\gamma_t(i) = prob(q_t = S_i \mid O, \lambda)$$

indicating the probability of being in state S_i at time slot t given the observation sequence O and model λ . The equation can be reformatted in terms of forward and backward variables:

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{prob(O \mid \lambda)} = \frac{\alpha_t(i)\beta_t(i)}{\sum_{i=1}^M \alpha_t(i)\beta_t(i)}$$

By maximizing $\gamma_t(i)$, the estimated state \hat{q}_t at time stamp t can be detected using $\hat{q}_t = \frac{\arg \max}{1 \le i \le M} [\gamma_t(i)], \ 1 \le t \le T$

The quantity \hat{q}_t is the estimated system state at the t-th instance of the state sequence of length T.

6.2 **Results and Discussion**

6.2.1 Experimental Methods

During an experiment, a subject was instructed to press a button whenever she or he swallows, and the smartphone shown in Figure 6-4 was used to record the breathing signal sensed by the wireless chest belt. A video camera was connected to a computer to record the movement of mouth and laryngopharynx during the experiment for validation purposes. The computer, smartphone, and button recorder were synchronized before each experiment session. The experiment setup is shown in Figure 6-4.



Figure 6-4: Experimental setup

Experiments were carried out on 6 subjects (2 female and 4 male) without any known swallow abnormalities. Each subject was asked to wear the instrumented chest-belt and have his or her lunch at his or her own pace. The lunch type was chosen by individual subjects based on their dietary preferences. It included diverse food types including rice, bread, salad, and cooked vegetarian and non-vegetarian items. Note that the subjects were allowed to drink during the experiments. However, since the results in this chapter are concentrated only for solid intake, the affected breathing cycles during drinking were first identified from video recording, and then removed during data processing.

6.2.2 Performance Indices

To evaluate the detection performance (i.e., both SVM-only and SVM followed by HMM), we adopted the metrics Precision and Recall, commonly used [27][53] in biomedical signal processing and information retrieval. Precision and Recall are defined as:

$$Precision = \frac{TP}{TP + FP} = \frac{Recognized \ swallows}{Retrieved \ swallows}$$

$$Recall = \frac{TP}{P} = \frac{Recognized \ swallows}{Relevant \ swallows}$$

In this definition, recognized swallows (i.e., true positives, TP) indicates the number of swallow events that are correctly detected. Retrieved swallows correspond to the number of detected swallows including both the TPs and the incorrectly detected swallows (i.e., false positives, FP). Relevant swallows (i.e., positive, P) refer to the number of actual swallow events annotated from video observations reflecting the ground truth.

6.2.3 Feature Extraction for Stage-1 Detection using SVM

As reported in our previous work [83], both time domain and frequency domain features can be used for detecting liquid swallows. The discriminative power of those feature types, however, can be different. Figure 6-5 shows the discrimination power of time (Figure 6-5:a) and frequency (Figure 6-5:b) domain features in solid swallow detection using SVM classifier. The merit of a feature in Figure 6-5:a and Figure 6-5:b refers to information gain [90], which is defined as the reduction in classification entropy (i.e., H(*)) with additional information provided by the corresponding feature about the target classes. Assuming A as the feature and C as the set of classes, the following two equations indicate class entropies before and after using the feature:

$$H(C) = -\sum_{c \in C} p(c) \log_2 p(c)$$
$$H(C \mid A) = -\sum_{a \in A} p(a) \sum_{c \in C} p(c \mid a) \log_2 p(c \mid a)$$



Figure 6-5: Feature discriminative property and ± 10 crossings as a classification

feature

A feature with higher merit indicates lower class entropy when this feature is adopted. Merit can be also used when feature reduction is needed in the presence of limited computational and storage resources.

For time domain features, as shown in Figure 6-5:a, where 128 sample points in normalized breathing cycles are used as features, sample points near the 27th and 53th sample points are more important than others in classification. For frequency domain features, as shown in Figure 6-5:b, where the first 64 FFT coefficients are used as features, lower frequency components have more discriminative power. It was also found that the discriminative power distribution of frequency domain features are more consistent across subjects, which is why the frequency-domain features were finally used in [83].

The second set of SVM classification features is derived from the first derivative of the breathing signal. As shown in Section 6.1, it was found that the swallows generally create more fluctuations in the breathing signal compared to the normal breathing cycles. To capture such fluctuations, an additional classification feature was derived from the first derivatives of the breathing signal. More specifically, the number of ± 10 crossings (NTC) is used as the feature, which is defined as the number of points in the breathing signal at which the first derivative of the signal is exactly +10 or -10. Compared to the number of zero crossings (NZC), NTC can better capture the swallow signatures.

Figure 6-5:c and Figure 6-5:d show an example comparison between a representative normal breathing cycle and a representative breathing cycle with swallow and their corresponding NZC and NTC of the first derivatives. Observe that while for both types of breathing cycles the NZC is 2, the number of ± 10 crossings (NTC) is 3 for the normal breathing and 5 for the breathing with swallow. The additional 2 NTCs (i.e., -10 crossings) are contributed

by the swallow event. Differences in NTCs were consistently observed between breathing cycles with and without swallows, thus indicating the usefulness of NTC of the first derivative as a useful classification feature for the SVM engine.

The third set of features is derived from the duration and amplitude of the breathing cycles before normalization. In summary, the SVM features used in this chapter include: first 5 Fourier transform coefficient, NTC, inhalation duration, exhalation duration, total breathing cycle duration, inhalation amplitude, and exhalation amplitude.

6.2.4 Swallow Detection with SVM

The features mentioned above were fed into the posterior probability SVM classifier described in Section 6.1.2. The classifier was trained and validated using data collected through experiments outlined in Figure 6-4. We have used a leave-one-out validation approach. Meaning a subject's data is excluded in training set if his or her data is used as the test set.

Figure 6-6 reports the distribution of the SVM-produced posterior probabilities (i.e., of probability of a cycle containing swallow signature) for breathing cycles with swallows and normal breathing cycles. The distribution was plotted from all classification data obtained during the experiment. In the absence of classification errors, there would have been only one bar at probability 1 for the cycles with swallows. Similarly, there would have been only one bar at probability 0 for the cycles without swallows. In Figure 6-6, it can be observed that in spite of some classification errors (i.e., indicated by the scattered bars over the probability axis) the SVM is able to separate the two cycle types fairly distinctly. Such errors are often caused by swallow signatures that are too short (in time) to be captured by the specified features, and by breathing cycle modulation by adjacent swallows [49].



Figure 6-6: Distribution of posterior probabilities with and without swallows

Applying a probability threshold Pth to the SVM-produced posterior probabilities it is possible to classify each cycle as a normal or swallow-containing breathing cycle in the following manner.

{normal breathing cycle, if $prob(1 | x_i) < Pth$ {breathing cycle with swallow, otherwise



Figure 6-7: Comparison between SVM-only and two-tier SVM+HMM mechanism Figure 6-7 shows the SVM-only classification accuracy (i.e., precision and recall) while the threshold Pth is changed in the range from 0.1 to 0.9. In Figure 6-7, each SVM-only performance point on the curve for a given subject corresponds to one threshold value and its
corresponding precision and recall performance pair. When the threshold varies from low to high, the precision increases, while the recall reduces, indicating fewer false positives and true positives. Since the breathing signal during lunch is imbalanced, meaning there are much more normal breathing cycles than breathing cycles with swallows, a higher threshold gives more preference over normal breathing cycles and therefore reduces the false positives, thus increasing the precision and decreasing the recall. The reverse effect was observed while lowering the detection threshold Pth. The SVM-only performance lines in Figure 6-7 provide a means for choosing an appropriate Pth for a required balance between precision and recall. Table 6-1 presents the precision and recall performance for all six subjects when an even threshold of 0.5 is chosen for swallow classification.

	SVM		SVM+HMM		
	Precision (%)	Recall(%)	Precision (%)	Recall(%)	
Subject 1	68	86	74	82	
Subject 2	67	100	74	96	
Subject 3	49	90	81	81	
Subject 4	54	98	72	93	
Subject 5	45	91	66	87	
Subject 6	80	80	83	74	

Table 6-1: Comparison between fixed threshold SVM-only and two-tier SVM+HMM mechanism

6.2.5 Improved Detection using HMM

Hidden Markov Model processing, as outlined in Section 6.1.4, was applied to the posterior probabilities output of SVM for improving detection performance. Such improvements are accomplished via correcting some of the SVM errors by the way of leveraging known

locality information in human swallow sequences. As described in Section 6.1.3, for each subject, both A (2x2) and B (2x10) matrices in the HMM model were computed based on experimental observations using video camera, push button, and data from the wireless chest-belt system. As an example, for Subject-1 the A matrix was found to be:

$$A = \begin{bmatrix} 0.76 & 0.24 \\ 0.96 & 0.04 \end{bmatrix}$$

and the B matrix was found to be:

$$B = \begin{bmatrix} 0.84 & 0.03 & 0.01 & 0 & 0.03 & 0.01 & 0.01 & 0.01 & 0.01 & 0.06 \\ 0.08 & 0.02 & 0 & 0 & 0.04 & 0 & 0.04 & 0.02 & 0.06 & 0.74 \end{bmatrix}$$

Swallow detection performance of the SVM+HMM approach is presented in Table 6-1 and in Figure 6-7. For each subject, there is one point for the SVM+HMM approach indicating the corresponding precision and recall performance. Observe that for subjects 1, 2, 3, and 6, the SVM+HMM point is situated higher and on the right in comparison to the line for the SVM-only approach. This indicates better performance of SVM+HMM compared to the SVM-only approach with all possible posterior probability thresholds. For the remaining two subjects (i.e., 4 and 5), the SVM+HMM performance point is on the SVM-only line, indicating that with certain posterior probability threshold, SVM-only can perform as good as the SVM+HMM approach. These results validate the overall usefulness of the proposed HMM processing by leveraging known swallow sequence locality information for removing certain classification errors that are introduced by the SVM-only approach.

6.3 Conclusion

In this chapter, we have presented a wireless and wearable solid food intake monitoring system. A novel support vector machine (SVM) and Hidden Markov Model (HMM) based

processing mechanism, which analyzes collected breathing signal and previously known swallow sequence locality information, was also presented. The system and processing mechanism was experimentally proven to be effective for solid food intake detection. Ongoing work on this topic includes: 1) developing an unsupervised swallow (both solid and liquid) detection mechanism for generalizability, 2) developing a detection and filtering mechanism for artifacts introduced by movement and speech, and 3) implementing a real-time swallow detection system that can be used by health researchers for retrieving dietary information from targeted population.

Chapter 7: Mealtime and Duration Monitoring

In this chapter, we present a wearable sensor system for estimating mealtime (i.e., time of the day for a meal) and meal duration based on those swallow-triggered apneas detected in breathing signals. Using two Respiratory Inductance Plethysmography (RIP) belts worn on the chest and abdomen as shown in Figure 7-1, swallow-triggered apneas are detected. Since the RIP belts do not rely on pressure or skin-contact to pick up breathing signal, they can be suitable for prolonged usage without minimal cosmetic and comfort issues.

Time and duration of meals were shown to be highly correlated with obesity. Ma *et al* [91] analyzed the mealtime of 499 people for 1 year and concluded that a greater number of eating episodes each day was associated with lower risk of obesity, whereas practices such as skipping breakfast was linked with higher obesity risk. Gluck *et al* [92] conducted experiments on 55 subjects and demonstrated that nighttime eaters, who consume considerable amount of food at night frequently tend to gain weight faster than non-nighttime eaters. Cleator [93] also showed that 52% of obese nighttime eaters reported normal weight before the onset of their nighttime eating habit. Andersen *et al* [94] carried out experiments lasted for 10 years, and reported that obese women with nighttime eating experienced a weight gain of 5.2kg on an average over 6 years.

Improper mealtime can have other negative impacts. Rogers *et al* [95] showed that subjects with night eating syndrome had less stage 2 and stage 3 sleep, which contributed to shorter total sleep time and lower sleep efficiency, and they are more likely to suffer from depression. Sassaroli *et al* [96] also found that night eating syndrome is highly correlated with anxiety.

In addition to the wearable belt, an accelerometer on the wrist of the dexterous hand has also been used to improve the swallow detection performance. The hand movement helps improving detection accuracy by supplying side information, so that when there is confusion due to uncertainties in detection using breathing signal, the hand movement information helps. The accelerometer is able to capture the characteristics of hand movement during eating, drinking, sitting, and talking artifacts.

In our previous work [97][98], we presented an early stage swallow detection system and algorithms for detecting liquid and solid food intake in the absence of artifacts. This chapter builds on those core system capabilities and leverages those algorithms in order to develop a system specifically for estimating the meal intake time and duration.

Specific contributions of this chapter include: 1) developing algorithmic solutions for handling artifacts including spontaneous swallows, talking, laughing, coughing and clearing throat, and 2) experimentally demonstrating the effectiveness of the proposed system and algorithms in a semi-controlled environment.

7.1 System Architecture

Figure 7-1 demonstrates the wearable sensor system, which is an extension of what was reported in [98]. The wearable system includes: 1) a pair of RIP belts and their associated control box (from zRIP Durabelt sum kit, Pro-Tech, Murrysville, PA) for collecting breathing signal, 2) a signal shaping circuit for amplifying and filtering the raw signal from the sensor to optimize signal-to-noise ratio (SNR) of ADC stage, 3) a μ Controller and Bluetooth subsystem equipped with 14-bit ADC channels and an 8-bit accelerometer for sampling and transmitting breathing and hand movement signal over Bluetooth to an external Android smart phone, and 4) a 3.7V 340mAh polymer rechargeable battery.



Figure 7-1: Components of the mealtime and duration monitoring system

The signal shaping circuit, μ Controller and Bluetooth subsystem, and battery are placed in a 4cm×2.5cm×2cm 3D-printed watch-like wrist unit worn on the wrist of the dexterous arm of a subject. The system is able to collect data continuously for over 20 hours on a single battery charge.

7.2 Processing Methods

Figure 7-2 depicts the overall logic and algorithmic architecture for the proposed mealtime and duration monitoring. The overall logic can be divided into three parts, namely, preprocessing, food intake detection, and meal intake analysis.

The preprocessing stage takes respiratory signal and wrist acceleration signal, and generates feature vectors. Each feature vector contains a number of features, which represent the unique characters of breathing in a cycle-by-cycle manner and of the wrist acceleration signal.

More specifically, features from breathing signal capture the signature of swallow apnea, whereas those from wrist acceleration represent hand movement during eating, drinking, sitting, talking, and other artifacts.

The food intake detection module utilizes a 3-stage hierarchical classifier, where the first classifier differentiates normal breathing cycles from non-normal breathing cycles (i.e., cycles with talking, solid swallow, and liquid swallow). The second classifier detects talking and swallowing (i.e., both solid and liquid swallows), and the third classifier differentiates between solid and liquid swallows.

The final architectural component (in Figure 7-2: The mealtime and duration detection scheme) is the meal intake analysis stage that estimates the time and duration of each meal intake episode based on the detected solid swallows from the previous food intake detection stage.

In the preprocessing stage, the respiratory signal from the from RIP sensors and the wrist acceleration signal, as described in Section 7.1, are sampled at 100Hz and fed into a digital low pass filter for removing quantization noise caused during the A-D conversion. The respiratory signal then passes through a breathing cycle extraction module, which performs a threshold based peak and valley detection operation as described in [85] on 33% overlapping 30-second windows.

The peak-valley detection algorithm selects peaks and valleys alternatively, such that the amplitude difference between each neighboring pair is larger than a threshold. The threshold is set to 0.3 (max $_{d(m)\in C} d(m) - \min_{d(n)\in C} d(n)$), where C is the data set that includes all the data points in the target 30-second window, and $\max_{d(m)\in C} d(m)$ and $\min_{d(n)\in C} d(n)$ are the amplitudes of the highest and the lowest sample points in the window respectively. In case of fluctuations below the threshold near a peak or valley, the highest or lowest points are selected. A threshold

of $(\max_{d(m)\in C} d(m) - \min_{d(n)\in C} d(n))$ is also applied in order to detect improper wear of the respiratory belts, in which case the received respiratory signal contains only low amplitude noise caused by its electronic components and the ADC



Figure 7-2: The mealtime and duration detection scheme

Each separated breathing cycle is normalized amplitude-wise and interpolated in time. The normalized and interpolated breathing cycles span between 0 and 100 amplitude-wise and have

128 sample points each. More specifically, the normalization of a cycle is performed for the valley-to-peak (inhale) and peak-to-valley (exhale) data segments individually. Each valley-to-peak and peak-to-valley data segment is normalized as follows:

$$normalized(i) = \frac{[data(i) - valley] \cdot 100}{(peak - valley)}$$

where data(i) corresponds to a data point in the valley-to-peak (inhale) or peak-to-valley (exhale) data segment, normalized(i) is the normalized data point corresponding to data(i), and peak and valley are the peak and valley points of the segment. Linear interpolation is adopted to interpolate each breathing cycle.

Considering the average length of a cycle of 3.77 seconds in our experiments, the normalized sampling rate after interpolation is 34Hz (128 sample points/3.77 seconds result in 34Hz). Although each cycle may have different time and amplitude ranges caused by various tidal volume [97] and respiratory frequency (i.e., variable person-to-person or cycle-to-cycle for the same person), the normalization process removes such variances, thus improving the generalizability of the proposed system in handling respiratory signals with various amplitude and frequency. The Feature Extraction module takes both the breathing signal and wrist acceleration data and generates 29 features from breathing signal and 12 features from wrist acceleration signal. Details about the features are provided in Section 7.3.3.

In the food intake detection phase, we used a 3-stage hierarchical classifier, which contains 3 individual classifiers: Classifier-1 for detecting normal breathing cycles and non-normal breathing cycles (including breathing cycles with talking, solid swallows and liquid swallows), Classifier-2 for identifying swallows (including both solid and liquid swallows) from breathing cycles with talking, and Classifier-3 for detecting solid and liquid swallows. For all three classifiers, Support Vector Machine (SVM) with posterior probability [99] are used. A posterior probability indicates the SVM-detected probability of a given breathing cycle to be either one of the two classes that the classifier is designed to differentiate. Details about the SVM with posterior probability are introduced in Section 6.1.2.

Hidden Markov Model (HMM) is applied to the posterior probability outputs of Classifier-1 and Classifier-2 in order to improve the detection performance by leveraging a-priori knowledge about any temporal locality present in the swallow and talking sequence. HMM is not applied to Classifier-3 (which differentiates solid and liquid swallows), because based on our experiments and observations, solid and liquid swallows do not demonstrate strong temporal locality. Meaning, swallows (i.e., solid or liquid) are generally not likely to happen in consecutive breathing cycles. Details about HMM modeling and its processing are discussed in Sections 6.1.3 and 6.1.4.

All 3 classifiers in the hierarchical classifier system use features extracted in the preprocessing module. Classifier-2 and Classifier-3 are triggered by the output of HMM-1 and HMM-2 respectively. Meaning Classifier-2 is deployed only when a breathing cycle is classified as non-normal (i.e., cycles with talking, solid swallows and liquid swallows), and Classifier-3 is applied when it is classified a swallow (i.e., both solid and liquid swallows).



Figure 7-3: Meal intake analysis algorithm

In our previous work [97][98], Support Vector Machine (SVM) was shown to be the most effective classifier for swallow detection. Like most commonly used SVM [77], the classification output for each breathing cycle is a class label, which is normal breathing cycle or breathing cycle with swallows. After analyzing the classification errors, it was found that many of those errors could be corrected by applying known temporal locality information in human swallow and breathing sequences. For example, people rarely swallow in many consecutive breathing cycles. Thus, whenever the classification output shows many consecutive breathing cycles with swallows, errors can be suspected and misclassified instances can be identified and removed by applying higher level techniques such as Hidden Markov Model (HMM). This motivates the cascading of SVM and HMM, as shown in Figure 7-2.

The meal intake analysis module takes the detected solid swallows from the food intake phase. Empirically, people execute solid swallows periodically during a meal. Therefore, when N (i.e., a threshold count) or more of solid swallows are detected in a window of M breathing cycles, those M cycles are categorized as the part of a meal intake episode. The detailed algorithm is illustrated in Figure 7-3. When a moving window is centered at a breathing cycle with solid swallows, and the number of solid swallows in the window exceeds the threshold N, the cycles among the solid swallows are classified as part of the meal intake episode.

7.3 Results

7.3.1 Experimental Methods

The experiments were carried out on 14 subjects (5 female and 9 male) without any known swallow abnormalities. The experiments were approved by Michigan State University's Institutional Review Board. Subjects were required to participate at least 3 sessions of experiments. The first 2 sessions were Type-1, and the last session was Type-2.



Figure 7-4: Experimental setup

During a Type-1 session, a subject was asked to have lunch first without talking, then drink water from a flask every 20 seconds for 10 times, and then rest for 10 minutes, and lastly converse with the experimenter freely for 10 minutes. A Type-2 session was the same as a Type-1 session except that the subject was allowed to talk when having lunch. Coughing, clearing throat, and other activities that impact breathing were also allowed and constantly observed during the whole experiment (both Type-1 and Type-2), and laughing was allowed during the 10-minute conversation of each experiment session. The food for lunch during the experiment included rice, bread, salad, fruit, cooked vegetarian and non-vegetarian items. Among the 14 subjects, 4 of them executed 3 Type-1 sessions and 1 Type-2 session lasted around 45 minutes, and in total, approximately 34 hours 30 minutes of data were collected.

The experimental setup is shown in Figure 7-4. During an experiment, a subject, wearing the instrumented system as described in Figure 7-1, was instructed to press a button whenever

she or he swallows during lunch or drinking, and the smartphone was used for recording the respiratory and hand movement signal captured by the system. A video camera was connected to a computer to record the movement of mouth and laryngopharynx during the experiment to indicate when the subject was talking and to validate the swallows. The computer, smartphone, and button recorder were synchronized before each session of experiment. The push-button and video information was used as a ground truth for verification purposes.

Note that the method of using a press button and video camera is more natural compared to the observer-based experiments as reported in [29][36], where an observer recorded the swallow events. Such observer-based method suffers from the observer-expectancy effect [100], meaning the expectation of experimenter affects the behavior of participants. In our experiment, we used a press button to indicate the ground truth, and a camera for verification purposes to avoid such observer-expectancy effect.

Normal breathing cycles, breathing cycle with talking, solid swallows, and liquid swallows were labelled correspondingly. Laughing, coughing, clearing throat and other artiflacts were labelled as talking. Data collected during the 10-minute rest were only used in the meal intake analysis phase as described in Section 7.3.5 for analyzing the impact of spontaneous swallows.

7.3.2 Performance Evaluation

To evaluate the performance of the SVM-only and the SVM-followed-by-HMM arrangements, we have used commonly used [29][32] biomedical signal processing and information retrieval performance indices, namely, Precision, Recall, and F-measure. They are defined as follows:

$$Precision = True Positives / (True Positives + False Positives)$$
(11)

$$Recall = True Positives / (True Positives + False Negatives)$$
(12)

$$F - measure = 2 \cdot Precision \cdot Recall/(Precision + Recall)$$
(13)

For Classifier-1 and HMM-1 (see Figure 7-2), which classify normal breathing cycles and non-normal breathing cycles (i.e., breathing cycle with talking, solid swallows and liquid swallows), True Positives indicate the non-normal breathing cycles that are correctly detected, False Positives means the normal breathing cycles that are mistakenly classified as non-normal breathing cycles, and False Negatives correspond to those non-normal breathing cycles that are wrongly detected as normal breathing cycles.

For Classifier-2 and HMM-2, which differentiate breathing cycles from talking and swallows (i.e., solid swallows and liquid swallows), True Positives means the swallows that are correctly detected, False Positives indicate those breathing cycles with talking that are mistakenly classified as swallows, and False Negatives are swallows detected as breathing cycles with talking.

For Classifier-3, which classifies solid and liquid swallows, True Positives depict the solid swallows that are corrected classified, False Positives indicate liquid swallows that are incorrectly detected as solid swallows, and False Negatives are solid swallows detected as liquid swallows.

F-measure is the harmonic mean of Precision and Recall, and it is used in the case of contradictory conclusions based on Precision and Recall individually when compared to the performance of two different classifiers. For example, when comparing two classifiers A and B, Precision of A is higher than that of B, whereas A's Recall is lower than B's, F-measure can then be used to draw a conclusion.

7.3.3 Feature Extraction

Category	Features				
	Based on non-	Duration of inhalation and exhalation (4)			
	normalized	Amplitude of inhalation and exhalation (4)			
	breathing	Breathing cycle duration and frequency (3)			
Respiratory signal related	cycles	Standard deviation of breathing signal in a cycle (1)			
		Number of local peaks (1)			
	Based on	Hist-60 (1)			
	normalized	rmalized Mean and standard deviation of breathing signal in a cycle (2)			
	breathing	First 10 FFT coefficients (10)			
	cycles	Energy of high frequency (>3 Hz) components (1)			
		± 10 crossing of first derivative(2)			
	Mean and stan	dard deviation of X and Y axis acceleration in the current breathing			
TT 1	cycle (4)				
Hand movement related	Mean and standard deviation of X and Y axis acceleration in the previous breathing				
	cycle (4)				
	Mean and standard deviation of X and Y axis acceleration in the breathing cycle				
	before previous one(4)				

Table 7-1: Features Extracted For Svm Classifiers

Table 7-1 demonstrates the 41 features extracted for SVM classifiers for food intake detection (numbers in the parenthesis indicate the number of features for each category). Note that the first 3 sets of features, i.e., duration of inhalation and exhalation, amplitudes of inhalation and exhalation, and breathing cycle duration, include both absolute and relative

numbers. For example, both absolute duration of inhalation and exhalation and the proportion of inhalation and exhalation as part of the whole breathing cycle are used as features. Absolute breathing cycle duration and the ratio between the duration and average duration of 2 neighboring breathing cycles are also used.

Hist-60 is derived as follows: 1) first divide the amplitude of the normalized breathing cycle into 10 equal intervals with ID 1 to 10, and then 2) set the number of samples falling into interval i as bin-i, i=1,2,...,10. Hist-60 is set to i when

$$\sum_{j=1}^{i} bin - j > 0.6 \cdot (total \, sample \, points) > \sum_{j=1}^{i-1} bin - j \tag{8}$$

 ± 10 crossing of first derivative is defined as the number of sample points whose first derivatives are either ± 10 or ± 10 crossing of first derivative has been used in [98] and proved to be effective in swallow detection.

7.3.4 Performance of Food Intake Detection

Food intake detection module as described in Figure 7-2 utilizes the features extracted from the preprocessing module, and detects normal breathing cycles, breathing cycles with talking, solid swallows, and liquid swallows. The food intake detection module uses the extracted features as shown in Table I from the preprocessing stage, and performs SVM and HMM classification. The detected normal breathing cycles, breathing cycles with talking, solid swallows, and liquid swallows are then fed in to the meal intake analysis module for meal intake episode detection.

SVM classifiers (i.e., Classifier-1, 2, and 3) can be used without involving the HMMs. As presented in Section 6.1.2, Classifier-1, 2 and 3 produce the posterior probability of each breathing cycle of being non-normal breathing cycle, swallow, and solid swallow respectively.

By comparing the posterior probability with a predefined threshold, detection results can be derived. For unbiased classifiers, which have no preference on any class, the threshold 0.5 can be adopted. For example, if the posterior probability produced by Classifier-1 is 0.3, the corresponding breathing cycle can be classified as a normal breathing cycle.

Figure 7-5 shows the performance measures as defined in Section 7.3.2 of this SVM-only food intake detection method with and without the hand movement features as illustrated in Table 7-1. Results are reported both for subject dependent and subject independent models.



■ Precision \square Recall \square F-measure

Figure 7-5:. Performance of SVM-only food intake detection method <u>Subject dependent model</u>: SVM classifiers were tested with data from one session and trained on the data from other sessions for each subject. <u>Subject independent model</u>: SVM classifiers were trained using the data collected from all subjects except the one whose data was used for testing.

Note that results shown in Figure 7-5 were based on the Type-1 experiments during which talking was not allowed during lunch. Figure 7-5:(a) shows the precision, recall and F-measure for Classifier-1, Classifier-2 and Classifier-3, when evaluated with subject dependent and without hand movement scenarios. Figure 7-5:(b) is for subject dependent and with hand movement scenarios. Finally, Figure 7-5:(c) is for subject independent and with hand movement case.

As expected, by providing useful side information, the optional hand movement feature did improve the overall detection performance. As shown in Figure **7-5**:(a) and (b), when the hand movement features are used, the performance of Classifier-2 improved significantly. However, the hand movement features have very limited impacts on the performance of Classifiers-1 and 3. The Subject-independent model has performance similar to the subject dependent model according to Figure 7-5:(b) and (c), which proves the good generalizability of the proposed method for food intake detection.

	SVM only		SVM+HMM			
Classifier	Precision	Recall	F-measure	Precision	Recall	F-measure
	(%)	(%)	(%)	(%)	(%)	(%)
1	86.6	76.6	81.3	85.1	81.7	83.4
2	60.4	80.4	69.0	72.8	71.3	72.0

Table 7-2: Comparison Between Svm-Only And Svm+Hmm Solutions

HMM modeling and processing, as introduced in Section 6.1.3 and 6.1.4, were applied on the posterior probability produced by the SVM classifiers to improve the performance by leveraging the temporal locality present in human swallowing and talking behavior. Table 7-2 demonstrates the comparison between SVM-only solution and SVM+HMM solution on Type-2 dataset, in which subjects can talk during lunch. The F-measure of SVM+HMM mechanism, which represents a harmonic mean of precision and recall, is constantly higher than that of the SVM-only solution.



Detected meal intake episode

T: Talking LS: Liquid Swallow SS: Solid Swallow NBC: Normal Breathing Cycle MIE: Meal Intake Episode Non-MIE: Non-Meal Intake Episode



7.3.5 Performance of Meal Intake Analysis

As the final step of mealtime and duration analysis, the meal intake analysis module as described in Figure 7-2, analyzes the detected solid swallows from food intake monitoring module and detects the meal intake episodes. Since people generally execute solid swallows periodically during a meal, a subject is considered to be within a meal when a more than a threshold number of solid swallows (i.e., N) are detected in a window of M breathing cycles. The detailed algorithm is depicted in Figure 7-3.



The optimum threshold N depends on the window size M. When M is large, the threshold also needs to be large. A large M causes the issue that small false positives may lead to large error in food intake episode detection, whereas a small M can lead to short detected food intake episodes for an actual meal. Figure 7-7 shows the optimum threshold selection for example window sizes of 100 seconds and 140 seconds. The error is defined as the difference in time between recorded lunch intake duration and the duration detected by the system. In the case of a window size of 100 seconds, the optimum threshold is 3 breathing cycles, whereas when the

window size is 140 seconds, the threshold needs to be adjusted to 4 breathing cycles. Note that the optimized M and N can be different for different subjects or even the type of food, but the analysis here provides a general guidance in selecting the parameters. In our experiments, window size of 100 seconds and threshold 3 were selected across all subjects and experiments.

Figure 7-6 shows an example of meal intake analysis with average food intake detection performance. Figure 7-6:(a) shows the results for Type-1 experiments, in which talking was not allowed during meal, and Figure 7-6:(b) depicts the results for Type-2 experiments, when talking was allowed. Figures on the top indicate the ground truth recorded by button pressing and the video camera. For instance, each solid swallow (SS) point during lunch corresponds to a solid swallow recorded by the button pressing. Similarly, each point during the drinking phase was a liquid swallow (LS), which was recorded by a button pressing, and the video record was used to verify the boundary between lunch and drinking and to indicate breathing cycles with talking. Detection results from food intake detection module are shown in the middle graphs in Figure 7-6. Each point in this section corresponds to a normal breathing cycle, breathing cycle with talking, solid swallow or liquid swallow detected by the food intake detection module as described in Figure 7-2. The graphs at the bottom of fig. 8 are the detected meal intake episodes by the meal intake analysis module in Figure 7-2. Each point in these graphs corresponds to a breathing cycle which either belongs to the meal intake episode (MIE) or not (Non-MIE). It can be seen that although there are quite a few detection errors in the food intake detection module, the meal intake analysis module can still detect the time and duration of meal intake episodes with fairly high accuracy.



More specifically, in Figure 7-6:(a), few breathing cycles are mistakenly classified as talking during meal. However, since the meal intake analysis module relies only on solid swallows, those few misdetections do not reduce the number of solid swallows in the window below the threshold N. As a result, the meal intake episode is still detected fairly accurately. Similarly, although few breathing cycles are erroneously classified as solid swallows during the talking session, the detected solid swallows are not enough (i.e., far below the threshold N) to be detected as meal intake episode. In Figure 7-6:(b), few of the solid swallows are not detected correctly at the beginning of the experiments, thus the detected meal intake episode is slightly shorter than the actual episode. Observe that many spontaneous swallows are detected during the resting session. At the beginning of talking session, three breathing cycles are erroneously classified as solid swallows, thus a short false positive is observed.

Figure 7-8 demonstrates the actual and detected duration of meal intake episode for both Type-1 and Type-2 experiments. It can be seen that the meal intake analysis module can generally estimate the meal intake episode for Type-1 experiments in an unbiased manner, but underestimates the episode for Type-2 experiments due to the fact that during lunch people sometimes converse for a while without feeding themselves.

7.4 Discussion

7.4.1 Restrictive Feature Selection

In the analysis in Section 7.3, we have used all the features listed in Table 7-1. While using all the features yields the best detection performance, it may not be desirable in situations where the available computational power is limited on devices such as smart phones, sensors, etc. In order to evaluate the proposed algorithms in such limited resource scenarios, the features can be ranked based on their discrimination power, which refers to information gain [90] defined as the reduction in classification entropy (i.e., H(*)) with additional information provided by the corresponding feature about the target classes. Assuming A as a feature and C as the set of target classes, the following equations indicate class entropy before and after using the feature:

$$H(C) = -\sum_{c \in C} p(c) \log_2 p(c)$$
⁽⁹⁾

$$H(C \mid A) = -\sum_{a \in A} p(a) \sum_{c \in C} p(c \mid a) \log_2 p(c \mid a)$$
(10)

A feature with high discrimination power brings low class entropy when the feature is utilized. This information gain-based feature selection algorithm has been used in [101] [102], and proven to be effective.



Figure 7-9: Performance of Classifier-1, 2 and 3 with different feature count

Figure 7-9 demonstrates the performance of Classifier-1, 2 and 3 when the number of adopted features is changed. Note that the HMM stage is bypassed in this analysis, and SVM [99] is used as the machine learning algorithm. Features are added into the feature set one by one based on their discrimination power evaluated based the mechanism suggested in [101] [102]. More specifically, the feature with highest discrimination power is first used, and then the feature with second highest discrimination power is added. The results in the figure represent the average performance for all the subjects using the Subject dependent model as described in Section 7.3.4.

As shown in Figure 7-9, performance of Classifier-1 stabilizes with 25 or more features, while performance of Classifier-2 stabilizes with 20 or more, and performance of Classifier-3 stabilizes with 10 or more features.

These results can be used to determine how many and which features should be used based on how much computational resources are available in the target platform such as a smart phone or sensor mote.

7.4.2 Benefits of Hierarchical Classifier

We have evaluated the performance difference between the proposed hierarchical classifier in Section 7.2 and a corresponding single-stage classifier with same classification objective. In the 3-stage hierarchical classifier, the first classifier detects normal breathing cycle and nonnormal breathing cycles (includes breathing cycles with talking, solid swallow and liquid swallow). The second classifier detects talking and swallowing (both solid and liquid swallow), and the third classifier identifies solid and liquid swallows.

A corresponding single stage classifier would classify normal breathing cycles, breathing cycles with talking, solid swallows, and liquid swallows. Such a classifier would use the features extracted by the Preprocessing module in Figure 7-2, and apply them to a single machine learning algorithm. We have used SVM for both the hierarchical classifier and the single stage classifier for their performance comparison as reported in Table 7-3.

	3-stage hierarchical classifier			Single classifier		
Classifier	Precision	D 11 (0/)	F-measure	Precision	D 11 (0/)	F-measure
	(%)	Recall (%)	(%)	(%)	Recall (%)	(%)
1	84.5	79.6	82.0	85.7	70.9	77.5
2	82.1	82.9	82.5	73.2	84.6	78.5
3	87.0	87.5	87.2	64.7	96.2	77.4

Table 7-3: Comparison of 3-Stage Hierarchical Classifier And Single Classifier

Performance of each classifier in the 3-stage hierarchical classification method is reported individually. For the single classifier, the performance is mapped to the corresponding stages of the hierarchical method. For example, performance of the single classifier compared to Classifier-1 (i.e., the one differentiates between normal breathing cycle and non-normal breathing cycle) in hierarchical method was computed by combining breathing cycles with talking, solid swallow and liquid swallow as a single category of non-normal breathing cycle. By comparing the F-measures of the two solutions, it can be observed that the proposed 3-stage hierarchical classifier constantly outperforms the single classifier solution, and justifies its additional complexity.

7.4.3 Performance of Existing Research

Passler *et al* [32] achieved 91.3% precision and 81.8% recall with an in-ear microphone, although only solid swallows were considered, and artifacts were excluded in their work. Amft *et al* [39] used inertial sensors to track the movement of arm and trunk, an ear microphone to record the food break down sound, and surface Electromyography (SEMG) electrodes and stethoscope microphone to detect swallowing activities. The derived detection of individual swallows resulted in 20% for precision and 68% for recall. Makeyev *et al* [34] deployed a throat microphone located over the laryngopharynx to detect swallow events, and the average accuracy is 66.7% for inter-subject model (cross-validation). In this paper, the cascading of Classifier-1 and Classifier-2 using the SVM+HMM scheme detects swallows, for which the combined precision and 60.6% for recall. Based on these observations, it can be concluded that the proposed detection algorithms in this paper offer a competitive mealtime estimation method with a brand new modality of sensing.

7.4.4 Spontaneous Swallows

Spontaneous swallows were identified and handled in this reported system. Spontaneous swallow is a protective aero-digestive reflex for airway protection, and it is caused by accumulated saliva and/or food remnants in the mouth [105]. For healthy subjects, the frequency of spontaneous swallows is about 1.22 times per minute [106].

In our experiments, the 10-minute rest within each experiment session was used to evaluate the performance for handling spontaneous swallows. As shown in Figure 7-6:(a), spontaneous swallows during resting were sometimes misdetected as talking, and as shown in Figure 7-6:(b), majority of those were detected as liquid swallows. However, as spontaneous swallows are less frequent and the meal intake analysis module only depends on solid swallows, spontaneous swallows did not impact on the detection performance of the of meal intake episodes, in general.

7.5 Conclusion

This paper presents a wearable sensor system for mealtime and duration detection based on breathing signal and hand movement analysis. Different from previous research, the proposed mechanism not only detects each individual swallows, but also detects the mealtime and duration information. Experiments were carried out on 14 subjects considering various artifacts that affect breathing signal, such as spontaneous swallows, talking, laughing, coughing, and clearing throat. Experimental results show that the proposed system and mechanism is an effective method for mealtime and duration monitoring.

Chapter 8: Proposed Work

In this thesis it was demonstrated that piezoelectric and RIP based sensor systems, and the proposed intake monitoring algorithms are able to detect swallow events, mealtime, and duration by observing a person's breathing signal. We first detect swallows by the way of detecting apneas extracted from breathing signal captured by a wearable wireless chest-belt. Afterwards, swallow pattern analysis is used for identifying swallows. Lastly, mealtime and duration detection is performed. Together with self-reporting at the high level of overall diet habits (i.e., the types of food and drinks etc.), the instrumented detection of swallow counts can offer an objective way to: 1) study the food and drink intake trends, and 2) estimate calorie intake. Building on the work done so far, we propose the following future search items.

8.1 Diet Volume Detection

To the best of our knowledge, most of the existing diet monitoring methods and algorithms focus on detecting the swallowing events. The evaluation metrics includes precision, recall and sensitivity. Although swallowing counts are generally correlated with volume of food consumed, the actual volume of bolus for each swallow may vary over time and across different subjects. Consequently, detecting swallow events without considering the volume of food intake may bring inaccuracy in estimating energy intake.

According to our experiments in Chapter 7, the volume for each swallow varies significantly across subjects, and also depends on the nature of the food. The following methods may be adopted for volume detection:

 Performing controlled experiment on subjects with various types of food to get the volume per swallow Using the swallow detection mechanism proposed to count the number of swallows.
 Together with the self-report of food type, the total volume of food can be estimated.

8.2 Choking and Coughing Detection

According to National Safety Council [107], choking is the third leading cause of home injury death for adults over 76, and second for people over 89. Choking is also fatal to children under 1 year old. Many factors can lead to choking, such as improperly chewed food, drinking alcohol, Parkinson's disease, etc.

Choking is the obstruction of the flow of air into lungs, and it prevents breathing. Prolonged choking time can result in asphyxia which leads to anoxia and is potentially fatal. Measures need to be taken within minutes before oxygen stored in the blood and lungs is depleted.

As shown in Chapter 3, holding breathe can cause constant output from the proposed chest belt systems. While choking, a victim may move irregularly causing motion effects. Therefore, by analyzing the component of breathing signal at different frequencies, choking can be detected.

Coughing is a protective reflex that clears the respiratory passages. It can be divided into three phases: (a) inhalation, (b) forced exhalation against closed glottis, and (c) violent release of air from lungs with opening of the glottis. Coughing is a common symptom of many diseases, such as virus and bacterial infection, respiratory tract infection, asthma, gastroesophageal reflux disease and etc.



Figure 8-1: Anticipated output of chest belts during coughing (a) one coughing within a breathing cycle (b) two consecutive coughs in a breathing cycle

Figure 8-1demonstrates the anticipated output of the chest belts during coughing based on the fact that coughing has three phases. If similar breathing signal can be observed during data collection, we should be able to detect coughing using matched filter based mechanism or machine learning mechanism.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] WHO, "WHO | Obesity and overweight," *WHO*. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs311/en/. [Accessed: 26-Jan-2013].
- [2] T. O. Cheng, "Fast food and obesity in China," J. Am. Coll. Cardiol., vol. 42, no. 4, pp. 773–773, Aug. 2003.
- [3] Yangfeng Wu, "Overweight and obesity in China," BMJ, vol. 333, Aug. 2006.
- [4] J. Chhatwal, M. Verma, and S. Riar, "Obesity among pre-adolescent and adolescents of a developing country (India).," *Asia Pac. J. Clin. Nutr.*, vol. 13, no. 3, pp. 231–235, 2004.
- [5] A. Misra, R. Pandey, J. Devi, R. Sharma, N. Vikram, and N. Khanna, "High prevalence of diabetes, obesity and dyslipidaemia in urban slum population in northern India.," *Int. J. Obes. Relat. Metab. Disord. J. Int. Assoc. Study Obes.*, vol. 25, no. 11, pp. 1722–1729, Nov. 2001.
- [6] R. H. Eckel and R. M. Krauss, "American Heart Association Call to Action: Obesity as a Major Risk Factor for Coronary Heart Disease," *Circulation*, vol. 97, no. 21, pp. 2099– 2100, Jun. 1998.
- [7] T. L. Visscher and J. C. Seidell, "The Public Health Impact of Obesity," Annu. Rev. Public Health, vol. 22, no. 1, pp. 355–375, 2001.
- [8] H. Jia and E. I. Lubetkin, "The impact of obesity on health-related quality-of-life in the general adult US population," *J. Public Health*, vol. 27, no. 2, pp. 156–164, Jun. 2005.
- [9] K. Weis, V. R. Taylor, and C. for D. C. and Prevention, *Measuring Healthy Days: Population Assessment of Health-Related Quality of Life.*.
- [10] L. S. Nielsen, K. V. Danielsen, and T. I. A. Sørensen, "Short sleep duration as a possible cause of obesity: critical analysis of the epidemiological evidence," *Obes. Rev.*, vol. 12, no. 2, pp. 78–92, 2011.
- [11] P. Björntorp, "Do stress reactions cause abdominal obesity and comorbidities?," Obes. Rev., vol. 2, no. 2, pp. 73–86, 2001.
- [12] A. Astrup, J. O. Hill, and S. Rössner, "The cause of obesity: are we barking up the wrong tree?," Obes. Rev., vol. 5, no. 3, pp. 125–127, 2004.
- [13] E. Jéquier, "Pathways to obesity," Int. J. Obes., vol. 26, no. Suppl2, pp. S12-S17, 2002.

- [14] V. A. Vance, S. J. Woodruff, L. J. McCargar, J. Husted, and R. M. Hanning, "Self-reported dietary energy intake of normal weight, overweight and obese adolescents," *Public Health Nutr.*, vol. 12, no. 2, pp. 222–227, Feb. 2009.
- [15] D. A. Schoeller, "Limitations in the assessment of dietary energy intake by self-report," *Metabolism*, vol. 44, Supplement 2, pp. 18–22, Feb. 1995.
- [16] C. D. Samuel-Hodge, L. M. Fernandez, C. F. Henríquez-Roldán, L. F. Johnston, and T. C. Keyserling, "A comparison of self-reported energy intake with total energy expenditure estimated by accelerometer and basal metabolic rate in African-American women with type 2 diabetes," *Diabetes Care*, vol. 27, no. 3, pp. 663–669, Mar. 2004.
- [17] B. Dong and S. Biswas, "Wearable networked sensing for human mobility and activity analytics: A systems study," in 2012 Fourth International Conference on Communication Systems and Networks (COMSNETS), 2012, pp. 1–6.
- [18] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher, "Activity recognition and monitoring using multiple sensors on different body positions," presented at the International Workshop on Wearable and Implantable Body Sensor Networks, 2006. BSN 2006, 2006, p. 4 pp. –116.
- [19] M. Sun and J. O. Hill, "A method for measuring mechanical work and work efficiency during human activities," *J. Biomech.*, vol. 26, no. 3, pp. 229–241, Mar. 1993.
- [20] S. E. Crouter, K. G. Clowers, and D. R. Bassett Jr, "A novel method for using accelerometer data to predict energy expenditure," J. Appl. Physiol. Bethesda Md 1985, vol. 100, no. 4, pp. 1324–1331, Apr. 2006.
- [21] U. Varshney, "Pervasive Healthcare and Wireless Health Monitoring," *Mob Netw Appl*, vol. 12, no. 2–3, pp. 113–127, Mar. 2007.
- [22] R. O. Dantas, M. K. Kern, B. T. Massey, W. J. Dodds, P. J. Kahrilas, J. G. Brasseur, I. J. Cook, and I. M. Lang, "Effect of swallowed bolus variables on oral and pharyngeal phases of swallowing," *Am. J. Physiol.*, vol. 258, no. 5 Pt 1, pp. G675–681, May 1990.
- [23] K. Zhang, F. X. Pi-Sunyer, and C. N. Boozer, "Improving energy expenditure estimation for physical activity," *Med. Sci. Sports Exerc.*, vol. 36, no. 5, pp. 883–889, May 2004.
- [24] J. Gates, G. G. Hartnell, and G. D. Gramigna, "Videofluoroscopy and Swallowing Studies for Neurologic Disease: A Primer1," *Radiographics*, vol. 26, no. 1, pp. e22–e22, Jan. 2006.
- [25] A. L. Perlman, P. M. Palmer, T. M. McCulloch, and D. J. Vandaele, "Electromyographic activity from human laryngeal, pharyngeal, and submental muscles during swallowing," J. *Appl. Physiol. Bethesda Md 1985*, vol. 86, no. 5, pp. 1663–1669, May 1999.
- [26] B. Dong and S. Biswas, "Swallow monitoring through apnea detection in breathing signal," presented at the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2012, pp. 6341–6344.

- [27] S. Passler and W.-J. Fischer, "Food Intake Activity Detection Using a Wearable Microphone System," in 2011 7th International Conference on Intelligent Environments (IE), 2011, pp. 298–301.
- [28] K. Takahashi, M. E. Groher, and K. Michi, "Methodology for detecting swallowing sounds," *Dysphagia*, vol. 9, no. 1, pp. 54–62, Dec. 1994.
- [29] O. Amft and G. Troster, "Methods for Detection and Classification of Normal Swallowing from Muscle Activation and Sound," presented at the Pervasive Health Conference and Workshops, 2006, 2006, pp. 1–10.
- [30] J. A. Cichero and B. E. Murdoch, "The physiologic cause of swallowing sounds: answers from heart sounds and vocal tract acoustics," *Dysphagia*, vol. 13, no. 1, pp. 39–52, 1998.
- [31] E. Keogh, S. Chu, D. Hart, and M. Pazzani, "An online algorithm for segmenting time series," in *ICDM 2001, Proceedings IEEE International Conference on Data Mining*, 2001, 2001, pp. 289–296.
- [32] S. Passler and W.-J. Fischer, "Food Intake Activity Detection Using a Wearable Microphone System," presented at the 2011 7th International Conference on Intelligent Environments (IE), 2011, pp. 298–301.
- [33] J. Nishimura and T. Kuroda, "Eating habits monitoring using wireless wearable in-ear microphone," in 3rd International Symposium on Wireless Pervasive Computing, 2008. ISWPC 2008, 2008, pp. 130–132.
- [34] O. Makeyev, P. Lopez-Meyer, S. Schuckers, W. Besio, and E. Sazonov, "Automatic food intake detection based on swallowing sounds," *Biomed. Signal Process. Control*, vol. 7, no. 6, pp. 649–656, Nov. 2012.
- [35] W. P. Walker and D. Bhatia, "Towards automated ingestion detection: swallow sounds," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2011, pp. 7075–7078, 2011.
- [36] E. Sazonov, S. Schuckers, P. Lopez-Meyer, O. Makeyev, N. Sazonova, E. L. Melanson, and M. Neuman, "Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior," *Physiol. Meas.*, vol. 29, no. 5, pp. 525–541, May 2008.
- [37] M. Aboofazeli and Z. Moussavi, "Analysis of swallowing sounds using hidden Markov models," *Med. Biol. Eng. Comput.*, vol. 46, no. 4, pp. 307–314, Apr. 2008.
- [38] O. Amft, H. Junker, and G. Troster, "Detection of eating and drinking arm gestures using inertial body-worn sensors," in *Ninth IEEE International Symposium on Wearable Computers*, 2005. Proceedings, 2005, pp. 160–163.
- [39] O. Amft and G. Tröster, "Recognition of dietary activity events using on-body sensors," *Artif. Intell. Med.*, vol. 42, no. 2, pp. 121–136, Feb. 2008.

- [40] L. Mioche, P. Bourdiol, J. F. Martin, and Y. Noël, "Variations in human masseter and temporalis muscle activity related to food texture during free and side-imposed mastication," *Arch. Oral Biol.*, vol. 44, no. 12, pp. 1005–1012, Dec. 1999.
- [41] C. S. Holger Nahrstaedt, "Swallow Detection Algorithm Based on Bioimpedance and EMG Measurements," pp. 91 – 96, 2012.
- [42] 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, Sep. 2012.
- [43] Alexandre Moreau–Gaudry, Abdelkebir Sabil, Gila Benchetrit, and Alain Franco, "Use of Respiratory Inductance Plethysmography for the Detection of Swallowing in the Elderly," *Dysphagia*, vol. 20, no. 4, pp. 297–302, Oct. 2005.
- [44] S. Damouras, E. Sejdic, C. M. Steele, and T. Chau, "An Online Swallow Detection Algorithm Based on the Quadratic Variation of Dual-Axis Accelerometry," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3352–3359, 2010.
- [45] E. Sejdic, C. M. Steele, and T. Chau, "Segmentation of Dual-Axis Swallowing Accelerometry Signals in Healthy Subjects With Analysis of Anthropometric Effects on Duration of Swallowing Activities," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1090– 1097, 2009.
- [46] A. Kandori, T. Yamamoto, Y. Sano, M. Oonuma, T. Miyashita, M. Murata, and S. Sakoda, "Simple Magnetic Swallowing Detection System," *IEEE Sens. J.*, vol. 12, no. 4, pp. 805– 811, 2012.
- [47] Y. Saeki and F. Takeda, "Proposal of Food Intake Measuring System in Medical Use and Its Discussion of Practical Capability," in *Knowledge-Based Intelligent Information and Engineering Systems*, R. Khosla, R. J. Howlett, and L. C. Jain, Eds. Springer Berlin Heidelberg, 2005, pp. 1266–1273.
- [48] K. Chang, S. Liu, H. Chu, J. Y. Hsu, C. Chen, T. Lin, C. Chen, and P. Huang, "The Diet-Aware Dining Table: Observing Dietary Behaviors over a Tabletop Surface," in *Pervasive Computing*, K. P. Fishkin, B. Schiele, P. Nixon, and A. Quigley, Eds. Springer Berlin Heidelberg, 2006, pp. 366–382.
- [49] B. Dong and S. Biswas, "Swallow monitoring through apnea detection in breathing signal," in 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2012, pp. 6341–6344.
- [50] J. E. Hall and A. C. Guyton, *Textbook of medical physiology*. Philadelphia, Pa.; London: Saunders, 2010.
- [51] G. Grimby, J. Bunn, and J. Mead, "Relative contribution of rib cage and abdomen to ventilation during exercise.," J. Appl. Physiol., vol. 24, no. 2, pp. 159–166, Feb. 1968.
- [52] K. Ashutosh, R. Gilbert, J. Auchincloss, and D. Peppi, "ASynchronous breathing movements in patients with chronic obstructive pulmonary disease.," *CHEST J.*, vol. 67, no. 5, pp. 553–557, May 1975.
- [53] O. Amft and G. Troster, "Methods for Detection and Classification of Normal Swallowing from Muscle Activation and Sound," in *Pervasive Health Conference and Workshops*, 2006, 2006, pp. 1–10.
- [54] B. J. Martin, J. A. Logemann, R. Shaker, and W. J. Dodds, "Coordination between respiration and swallowing: respiratory phase relationships and temporal integration," J. *Appl. Physiol. Bethesda Md 1985*, vol. 76, no. 2, pp. 714–723, Feb. 1994.
- [55] M. S. Klahn and A. L. Perlman, "Temporal and durational patterns associating respiration and swallowing," *Dysphagia*, vol. 14, no. 3, pp. 131–138, 1999.
- [56] H. G. Preiksaitis and C. A. Mills, "Coordination of breathing and swallowing: effects of bolus consistency and presentation in normal adults," *J. Appl. Physiol. Bethesda Md 1985*, vol. 81, no. 4, pp. 1707–1714, Oct. 1996.
- [57] H. Nilsson, O. Ekberg, R. Olsson, O. Kjellin, and B. Hindfelt, "Quantitative assessment of swallowing in healthy adults," *Dysphagia*, vol. 11, no. 2, pp. 110–116, 1996.
- [58] H. G. Dietz, "Pneumatic breathing belt sensor with minimum space maintaining tapes," 4602643, 29-Jul-1986.
- [59] P. Corbishley and E. Rodriguez-Villegas, "Breathing Detection: Towards a Miniaturized, Wearable, Battery-Operated Monitoring System," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 1, pp. 196–204, 2008.
- [60] M. Y.-W. Chia, S. W. Leong, C. K. Sim, and K. M. Chan, "Through-wall UWB radar operating within FCC's mask for sensing heart beat and breathing rate," in *Microwave Conference, 2005 European*, 2005, vol. 3, p. 4 pp.–.
- [61] K. Mukai, Y. Yonezawa, H. Ogawa, H. Maki, and W. M. Caldwell, "A remote monitor of bed patient cardiac vibration, respiration and movement," *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Conf.*, vol. 2009, pp. 5191–5194, 2009.
- [62] S. E. Crouter, K. G. Clowers, and D. R. Bassett Jr, "A novel method for using accelerometer data to predict energy expenditure," J. Appl. Physiol. Bethesda Md 1985, vol. 100, no. 4, pp. 1324–1331, Apr. 2006.
- [63] A. M. Swartz, S. J. Strath, D. R. Bassett Jr, W. L. O'Brien, G. A. King, and B. E. Ainsworth, "Estimation of energy expenditure using CSA accelerometers at hip and wrist sites," *Med. Sci. Sports Exerc.*, vol. 32, no. 9 Suppl, pp. S450–456, Sep. 2000.

- [64] M. H. Jones, R. Goubran, and F. Knoefel, "Reliable respiratory rate estimation from a bed pressure array," Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Conf., vol. 1, pp. 6410–6413, 2006.
- [65] M. Holtzman, D. Townsend, R. Goubran, and F. Knoefel, "Breathing sensor selection during movement," Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Conf., vol. 2011, pp. 381–384, 2011.
- [66] A. Bates, M. Ling, C. Geng, A. Turk, and D. K. Arvind, "Accelerometer-Based Respiratory Measurement During Speech," in 2011 International Conference on Body Sensor Networks (BSN), 2011, pp. 95–100.
- [67] A. Jin, B. Yin, G. Morren, H. Duric, and R. M. Aarts, "Performance evaluation of a tri-axial accelerometry-based respiration monitoring for ambient assisted living," *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Conf.*, vol. 2009, pp. 5677–5680, 2009.
- [68] P. D. Hung, S. Bonnet, R. Guillemaud, E. Castelli, and P. T. N. Yen, "Estimation of respiratory waveform using an accelerometer," in 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008, 2008, pp. 1493–1496.
- [69] M. H. T. Reinvuo, "Measurement of respiratory rate with high-resolution accelerometer and emfit pressure sensor," pp. 192 – 195, 2006.
- [70] B. B. Koo, C. Drummond, S. Surovec, N. Johnson, S. A. Marvin, and S. Redline, "Validation of a Polyvinylidene Fluoride Impedance Sensor for Respiratory Event Classification during Polysomnography," J. Clin. Sleep Med., Oct. 2011.
- [71] K. P. Sau, D. Khastgir, and T. K. Chaki, "Electrical conductivity of carbon black and carbon fibre filled silicone rubber composites," *Angew. Makromol. Chem.*, vol. 258, no. 1, pp. 11–17, 1998.
- [72] G. Gautschi, Piezoelectric Sensorics: Force, Strain, Pressure, Acceleration and Acoustic Emission Sensors, Materials and Amplifiers. Springer, 2002.
- [73] G. G. Mazeika, "Respiratory Inductance Plethysmography An Introduction," 2007.
- [74] R. O. Dantas, M. K. Kern, B. T. Massey, W. J. Dodds, P. J. Kahrilas, J. G. Brasseur, I. J. Cook, and I. M. Lang, "Effect of swallowed bolus variables on oral and pharyngeal phases of swallowing," *Am. J. Physiol.*, vol. 258, no. 5 Pt 1, pp. G675–681, May 1990.
- [75] G. Turin, "An introduction to matched filters," *IRE Trans. Inf. Theory*, vol. 6, no. 3, pp. 311 –329, Jun. 1960.
- [76] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *SIGKDD Explor Newsl*, vol. 11, no. 1, pp. 10–18, Nov. 2009.

- [77] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. Wiley-Interscience, 2000.
- [78] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing (3rd Edition)*, 3rd ed. Prentice Hall, 2009.
- [79] M. A. Hall and G. Holmes, "Benchmarking attribute selection techniques for discrete class data mining," *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 6, pp. 1437–1447, 2003.
- [80] L. Tarassenko, A Guide to Neural Computing Applications. John Wiley & Sons, 1998.
- [81] A. K. Jain, J. Mao, and K. M. Mohiuddin, "Artificial neural networks: a tutorial," *Computer*, vol. 29, no. 3, pp. 31–44, 1996.
- [82] L. Rokach, *Data Mining with Decision Trees: Theory and Applications*. World Scientific, 2008.
- [83] Bo Dong, Subir Biswas, "Noninvasive wearable diet monitoring through breathing signal analysis," presented at the SPIE Defense, Security, and Sensing, Baltimore, Maryland, United States, 2013.
- [84] Bo Dong and Subir Biswas, "Liquid Intake Monitoring Through Breathing Signal Using Machine Learning," in SPIE Defense, Security, and Sensing, Baltimore, Maryland, United States, 2013.
- [85] A. J. Wilson, C. I. Franks, and I. L. Freeston, "Algorithms for the detection of breaths from respiratory waveform recordings of infants," *Med. Biol. Eng. Comput.*, vol. 20, no. 3, pp. 286–292, May 1982.
- [86] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, 1992, pp. 144–152.
- [87] G. Wahba and G. Wahba, Support Vector Machines, Reproducing Kernel Hilbert Spaces and the Randomized GACV. 1998.
- [88] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [89] L. E. Baum and J. A. Eagon, "An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology," *Bull. Am. Math. Soc.*, vol. 73, no. 3, pp. 360–363, 1967.
- [90] M. A. Hall and G. Holmes, "Benchmarking attribute selection techniques for discrete class data mining," *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 6, pp. 1437–1447, 2003.

- [91] Y. Ma, E. R. Bertone, E. J. Stanek, G. W. Reed, J. R. Hebert, N. L. Cohen, P. A. Merriam, and I. S. Ockene, "Association between Eating Patterns and Obesity in a Free-living US Adult Population," *Am. J. Epidemiol.*, vol. 158, no. 1, pp. 85–92, Jul. 2003.
- [92] M. E. Gluck, C. A. Venti, A. D. Salbe, and J. Krakoff, "Nighttime eating: commonly observed and related to weight gain in an inpatient food intake study," *Am. J. Clin. Nutr.*, vol. 88, no. 4, pp. 900–905, Oct. 2008.
- [93] J. Cleator, J. Abbott, P. Judd, C. Sutton, and J. P. H. Wilding, "Night eating syndrome: implications for severe obesity," *Nutr. Diabetes*, vol. 2, no. 9, p. e44, Sep. 2012.
- [94] G. S. Andersen, A. J. Stunkard, T. I. A. Sørensen, L. Petersen, and B. L. Heitmann, "Night eating and weight change in middle-aged men and women," *Int. J. Obes. Relat. Metab. Disord. J. Int. Assoc. Study Obes.*, vol. 28, no. 10, pp. 1338–1343, Oct. 2004.
- [95] N. L. Rogers, D. F. Dinges, K. C. Allison, G. Maislin, N. Martino, J. P. O'Reardon, and A. J. Stunkard, "Assessment of sleep in women with night eating syndrome," *Sleep*, vol. 29, no. 6, pp. 814–819, Jun. 2006.
- [96] S. Sassaroli, G. M. Ruggiero, P. Vinai, S. Cardetti, G. Carpegna, N. Ferrato, P. Vallauri, D. Masante, S. Scarone, S. Bertelli, R. Bidone, L. Busetto, and S. Sampietro, "Daily and nightly anxiety among patients affected by night eating syndrome and binge eating disorder," *Eat. Disord.*, vol. 17, no. 2, pp. 140–145, Apr. 2009.
- [97] Bo Dong, Subir Biswas, "Wearable Diet Monitoring through Breathing Signal Analysis," presented at the 2013 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 2013.
- [98] B. Dong, S. Biswas, R. Gernhardt, and J. Schlemminger, "A Mobile Food Intake Monitoring System Based on Breathing Signal Analysis," in *Proceedings of the 8th International Conference on Body Area Networks*, ICST, Brussels, Belgium, Belgium, 2013, pp. 165–168.
- [99] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, New York, NY, USA, 1992, pp. 144–152.
- [100] D. F. Balph and M. H. Balph, "On the Psychology of Watching Birds: The Problem of Observer-Expectancy Bias," *The Auk*, vol. 100, no. 3, pp. 755–757, Jul. 1983.
- [101] H. Liu, J. Sun, L. Liu, and H. Zhang, "Feature selection with dynamic mutual information," *Pattern Recognit.*, vol. 42, no. 7, pp. 1330–1339, Jul. 2009.
- [102] C. S. Dhir, N. Iqbal, and S.-Y. Lee, "Efficient feature selection based on information gain criterion for face recognition," in *International Conference on Information Acquisition*, 2007. ICLA '07, 2007, pp. 523–527.

- [103] H. B. Borges, C. N. Silla Jr., and J. C. Nievola, "An evaluation of global-model hierarchical classification algorithms for hierarchical classification problems with single path of labels," *Comput. Math. Appl.*, vol. 66, no. 10, pp. 1991–2002, Dec. 2013.
- [104] E. P. Costa, D. C. D. Computação, I. S. Carlos, A. C. Lorena, and I. S. Carlos, A Review of *Performance Evaluation Measures for Hierarchical Classifiers*.
- [105] M. A. Crary, G. D. Carnaby, I. Sia, A. Khanna, and M. F. Waters, "Spontaneous swallowing frequency has potential to identify dysphagia in acute stroke," *Stroke J. Cereb. Circ.*, vol. 44, no. 12, pp. 3452–3457, Dec. 2013.
- [106] M. Pehlivan, N. Yüceyar, C. Ertekin, G. Çelebi, M. Ertaş, T. Kalayci, and I. Aydoğdu, "An electronic device measuring the frequency of spontaneous swallowing: Digital Phagometer," *Dysphagia*, vol. 11, no. 4, pp. 259–264, Sep. 1996.

[107] [Online].

Available:

http://www.nsc.org/safety_home/HomeandRecreationalSafety/Pages/Choking.aspx. [Accessed: 26-Nov-2013].