Wearable Long-term Social Sensing for Mental Wellbeing

Long Jiang[1], Bin Gao[1], Jun Gu[1], Yuanpeng Chen[1], Zhao Gao[2,3], Xiaole Ma[2], Keith M. Kendrick[2], W.L. Woo[4]
[1]School of Automation Engineering, [2]The Clinical Hospital of Chengdu Brain Science Institute, MOE Key Laboratory for Neuroinformation, [3]School of Foreign Languages, University of Electronic Science and Technology of China, China
[4]School of Electrical and Electronic Engineering, Newcastle University, England, UK.

Abstract — Long-term wellbeing monitoring is an underlying theme in many local and national policies and procedures outlined by governments and health care services. Natural, efficacious, and trustworthy monitoring by using wearable sensors is necessary for researchers to find and establish the interrelationships of Affective Computing (AC), Social Signal Processing (SSP) and Physical Mental Health (PMH). This paper investigated how technology can help to objectively monitor an individual's wellbeing in a naturalistic environment. For this purpose, we designed and implemented a wearable device with the integration of multi-sensors which consist of audio sensing, behavior monitoring, environment and physiological sensing. In order to avoid privacy issues, four audio-wellbeing features are embedded into a wearable hardware platform to automatically evaluate speech information without preserving raw audio data. In addition, four weeks of long-term monitoring experiment studies have been conducted in conjunction with a series of wellbeing questionnaires in a group of students to investigate objectively relationships between physical and mental health by utilizing data from speech, behavioral activities and ambient factors in a completely natural daily situation.

Index Terms — Long-term monitoring, Audio and activity features, Wearable device

I. INTRODUCTION

BEING socially engaged is human nature, whether it is with a single important person in one's life or a large group of friends. The current methods of measuring an individual's mental wellbeing, and how it may change over time, commonly involve the use of subjective rating scales. Most scales impose a temporal restriction on the assessment period. In order to obtain objective assessment, there has been increasing interest in the interrelationships of Affective Computing (AC) [1], Social Signal Processing (SSP) [2] and Physical Mental Health (PMH) [3]. These subjects have attracted researchers from various fields such as psychology, healthcare, ambient intelligence and computer science [4-6].

A variety of wearable sensors have been used to evaluate both physical activity and social interaction in practical ways. These include location beacons, accelerometers, cameras, acoustic sensors, light sensors, and physiological sensors [7]. Biosensors are capable of measuring significant physiological parameters such as heart rate, blood pressure, body and skin temperature, oxygen saturation, respiration rate, electrocardiogram etc. [8]. In addition, the fusion technology of hybridizing light, temperature, gas, microphone, accelerometers, a pressure sensor, and an infrared (IR) sensor for achieving more robust motion detection [9-10] have been investigated. In [11], a mobile sensing platform using an embedding microphone, visible light, pho-transistor, 3-axis digital accelerometer, digital barometer, temperature etc. to evaluate complex human activities has been described. In [12], a wearable computing platform (Electronic Badges) has been implemented for measuring and analyzing human behavior to recognize common daily activities and extracting speech features in real time to capture nonlinguistic social signals. Wearable sensing can provide revolutionary, objective information regarding the behavior and social signals produced by people in their natural environment [13].

Notwithstanding the above, wearable sensing enables diagnostic capabilities which include physiological and biochemical sensing, smartphone sensing, as well as motion sensing [14-16]. The application areas are numerous and mainly include Autism Spectrum Disorders (ASD) [17], Care for the Elderly, Behavioral therapy [18], stress and mental health monitoring etc. Monitoring of ASD has been conducted by applying activity tracking, stress notification/monitoring, geo-fencing (creating safe-zones), emergency response, discrete vibration alert and the capacity to add unlimited contacts [19]. Elderly Monitoring of the Elderly has been conducted using wearable sensor devices, localization, and activity recognition techniques which can support the caregiver in obtaining information on their daily life as well as evaluating their physical and mental health status [20-21]. Big data systems for healthcare have been established with distributed wearable sensors which enable assessment of the global elderly population globally by continuous monitoring and acquiring and analyzing data from the distributed devices [22-23]. Behavioral therapy requires continuous long-term physiological monitoring for the treatment of conditions such as drug addiction [24-25], obesity issues [26] and alcohol addiction [27]. In [24], the wearable sensors consist of a neoprene band that contains circuitry for measuring electrodermal activity (EDA), 3-axis motion, temperature and electrocardiogram (ECG). This can provide an additional therapeutic channel for patients with drug-addiction and post-traumatic stress disorder. There are huge opportunities for the use of wearable sensing in assessing mental health, such as mood, anxiety, emotions, depression and stress [28-29], where continuous physiological monitoring would provide a vital complement to clinical visits [30]. In [31], a headband-based mental-health monitoring system is proposed by using electroencephalography (EEG), together with a chest-band heart-rate variability (HRV) and also a skin-conductance (SC) sensor. In [32], researchers proposed a wearable eye tracking device as a novel measurement technique for day-to-day mental health monitoring. In [33], researchers proposed a novel personalized system comprised of a wearable monitoring system with embedded sensors and a smartphone to achieve better management of patients affected by mental disorders such as bipolar disorder. In [34-42], Smartphone apps or wearable sensor suites are used to collect continuous measurements indexing stress, emotion and mental illness in the natural environment. In addition, researchers have proposed sensor-enabled smartphone apps such as StressSense [34], BeWell+ [29] to recognize stress from the human voice and a wearable sensor suite such as Affective Q-Sensor [38], bio-monitoring devices [37, 40-41] and AutoSense [42] to monitor physiological and mental states.

In recent years, there has been a growing concern about mental health issues on college campuses [43]. There have been a number of studies on mental health issues among college students using wearable sensor, mobile phone and survey approaches [43-46]. In [45], researchers have developed a machine learning algorithm to distinguish between happy and unhappy college students. This algorithm evaluates the relationship between different components of wellbeing including happiness, health, energy, alertness and stress. However, there exist several problems in mental health monitoring. One of the more serious problems is privacy-protection [47] where analysis of audio signals for social activities and nonlinguistic signals is forbidden (body language, tone of voice, facial expression, ambient factors) [48-50]. In addition, other challenges include long-term monitoring, real-time data computing and analysis, portability and efficient evaluation.

Thus, in this paper, we develop a long-term wearable wellbeing sensing watch for unobtrusively and continuously capturing the granular details of behaviors and contexts which might provide important cues about onset of anxiety and autism. In particular, a set of integrated wellbeing sensing features are benchmarked to objectively recognize mental issues. Specifically, we tested adult-autistic spectrum trait levels in two groups of students divided into those with high scores or low scores (on the Autism Spectrum Quotient – ASQ). Subjects were monitored continuously for four weeks of data collection to explore all sensing features with indices related to ASQ score. In addition, in order to protect privacy, we have embedded the audio wellbeing feature processing in the hardware to avoid directly recording their voices and finally the correlation between wearable wellbeing sensing features and adult-autistic traits have been established.

The paper is organized as follows. Section II introduces the proposed wearable sensing architecture, implementation of feature extraction algorithms in wearable hardware and the implementation of the experimental set up. Experimental analysis and a series of correlation studies with adult-autistic trait levels are presented in Section III. Discussion and analysis based on the most relevant features are presented in Section IV. Section V presents the conclusions from the paper.

II. Methodology And Experiment

A. *Wearable social-mental-health sensing platform*

In this section, the proposed wearable social-mental-health sensing architecture is specified. The system is constructed taking into account social awareness so as to unobtrusively and continuously collect acoustic, physical activity and physiological indices in a completely natural and unpredictable situation. The wearable device consists of an ARM-Cortex4 microcontroller with DSP function for audio feature calculation, a variety of digital sensors for collecting multi-modal data from the environment, physiological signals and behavioral activity. In addition, specific components [5] consist of a microSD card for long-term data storage, a power management unit with a 2200mAh lithium battery, OLED screen, a debug port for download and debug firmware program and a USB port. The block diagram of the wearable platform is shown in Figure 1.

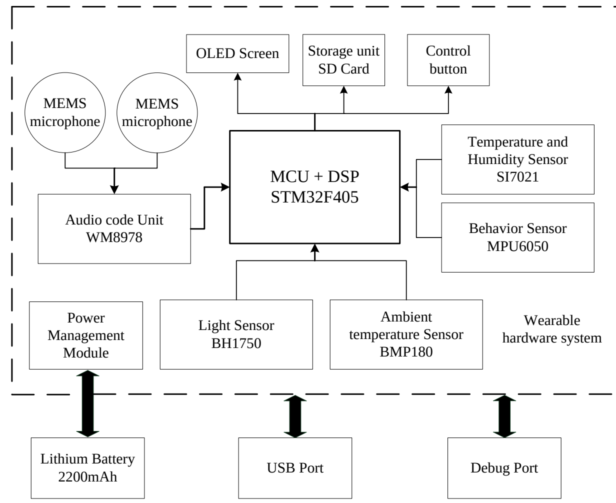


Figure 1. System block diagram of wearable device

In order to protect personal privacy, the wearable device extracts social related audio features and deletes raw data automatically. We embed four social-related audio features into the wearable hardware platform: Energy, Entropy, Brightness [5] and Formant [12]. We use ADC (analog digital converter) sampling and filtering for the digital signal of the audio code unit WM8978. In addition, STM32F405 of I2C (Inter Integrated-Circuit) bus protocol with WM8978 to control amplification gain of audio signal filtered is employed. Specifically, eight thread-based controlling work packages are implemented. These are an MPU-thread for collecting MPU6050 sensor data, a sensor-thread for collecting data from BH1750, BMP180 and SI7021 sensors, an audio-thread for collecting raw voice data, a DSP-thread for extracting audio features, a file-thread for storing sensor data and audio features, a battery-thread and oled-thread for charging management and data display, respectively.

The implementation diagram of the firmware is shown in Figure 2. The double buffer caching mechanism is developed in order to enhance the recording stage and reduce power consumption. Data from the sensors and microphones are stored every 10 minutes both separately and simultaneously by using embedded system (RT-Thread) and file system (FAT-32). The raw voice data are stored temporarily and the file-thread sends a signal into the DSP-thread for extracting audio features. Thus, in order to protect personal privacy during long-time health monitoring, the file thread will automatically delete the saved voice file after analysis and only retain the audio features. Specifically, because it is a multi-tasking operating system, the old data processing and the acquisition of the new data can be processed in ‘parallel’. Although this is pseudo-parallel process, the delay between new data collection and old data processing is T (10 minutes)).

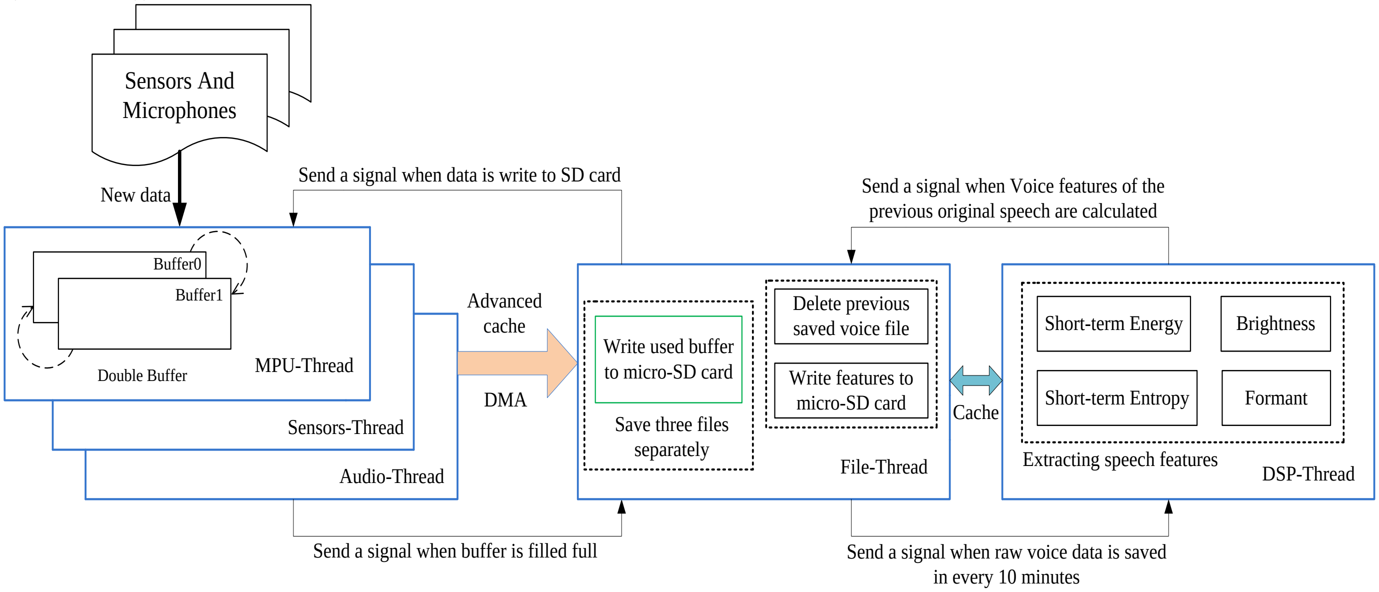


Figure 2. Implementation framework of the firmware

In the proposed wearable platform, the sampling rate of the activity sensor (MPU6050), environment sensors (BH1750, BMP180) and body sensor (SI7021) are 100Hz, 0.1Hz and 0.1Hz, respectively. The sampling rate of the audio sensor is 8KHz.

B. Embedded audio social features

The audio social features extraction algorithms embedded which are based on floating-point computing unit and DSP software library of STM32F405 is described as following:

1) The original voice data are divided into *M* segments which each segment is approximately 30ms. The short-term energy calculating formula for the *i*-th frame audio signal is expressed as:

$$E(i)=\sum_{n=0}^{L-1} y_i^2(n), \qquad 1\le i\le fn \tag{1}$$

where L is frame length, fn is the total number of frames after framing, $E(i)$ is i-th frame energy, $y_i(n)$ is the speech signal after using Hamming window for preprocessing.

2) The spectral entropy calculation method firstly calculates the spectral probability density, and then uses the spectral probability density to calculate the short-term spectral entropy of a frame for the original audio signal.

Assuming that the time-domain waveform of the noise-containing signal is $x(n)$, the i-th frame speech signal obtained by Hamming window is $x_i(m)$, the energy spectrum of the k-th spectral frequency component $f(k)$ is expressed as $Y_i(k)$ by applying FFT (Fast Fourier Transform) of $x_i(m)$. The normalized spectral probability density function for each frequency component can be obtained, namely

$$p_i(k)=\frac{Y_i(k)}{\sum_{l=0}^{N/2} Y_i(l)} \tag{2}$$

where $p_i(k)$ is the probability density of the k-th frequency component $f(k)$ corresponding to the i-th frame, N is the length of a FFT window.

The short-time entropy spectrum of each analyzed audio frame is expressed as:

$$H_{\mathrm{i}}=-\sum_{k=0}^{N/2} p_i(k)\log p_i(k) \tag{3}$$

where H_{i} is short-term spectral entropy calculated of the original data of the i-th frame.

3) The brightness of audio data is a phonetic feature with a high correlation with emotion. The essence is the centroid of energy spectrum. The specific calculation formula is defined as follows:

$$Brightness=\frac{\int_0^{w_0} w\,|F(w)|^2\,dw}{\int_0^{w_0} |F(w)|^2\,dw} \tag{4}$$

where w is the frequency corresponding to the spectral component, $F(w)$ is magnitude of a frequency point.

The discrete calculation formula of the brightness can be expressed as:

$$B_i=\frac{\sum_{k=0}^{N/2} w_k\cdot|y_k|^2}{\sum_{k=0}^{N/2}|y_k|^2} \tag{5}$$

where B_i is i-th frame brightness, w_k is the frequency of the k-th spectrum line, y_k is the value of energy spectrum of the k-th spectrum line.

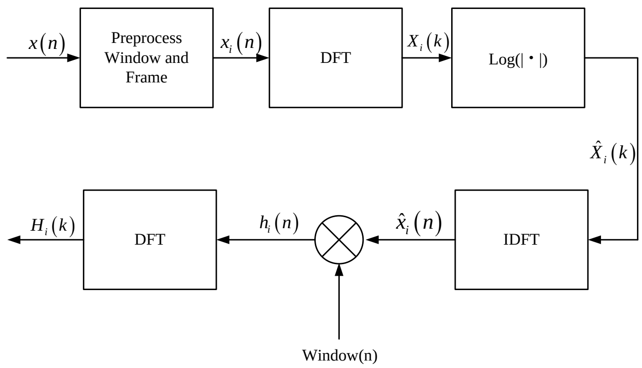
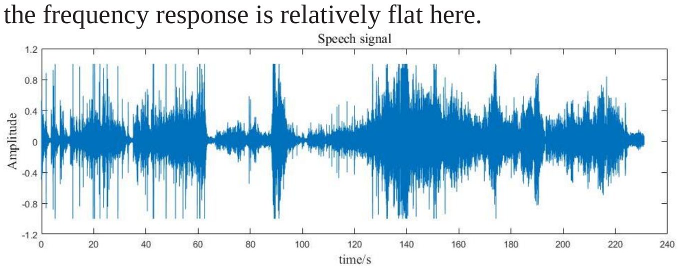


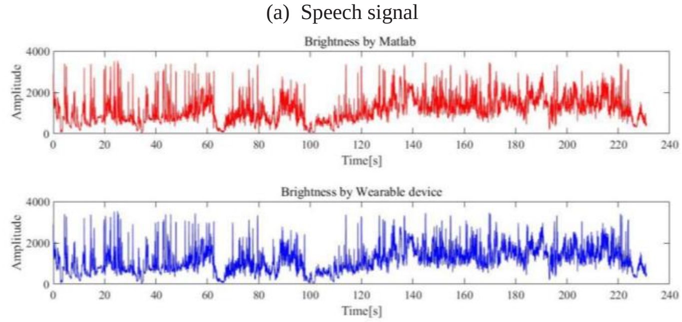
Figure 3. Flow chart to calculate Envelope of speech signal

4) We used the cepstrum method to calculate the short-time formant. Figure 3 shows the workflow to calculate the envelope $H_i(k)$ of a frame. Each formant parameter is the maximum point of the envelope calculated. We have calculated up to 5 formants.

In order to verify the correctness of the calculated speech features on wearable hardware, we collected a period of more than 4 minutes of speech signals. Both offline MATLAB calculation and wearable device processing of the four audio features have been conducted for comparison. Figure 4 illustrates the result of the brightness calculated by MATLAB and the wearable device. As can be seen, the brightness computed by the hardware is similar to that by MATLAB. Similar performance is also reflected in the energy and entropy calculations. However, during the calculation of the formant, the hardware system used a DSP library which may sporadically lose a formant frequency point as it is difficult to identify a formant within the frequency range from 1250Hz to 1600Hz as the frequency response is relatively flat here.



(a) Speech signal



(b) Brightness calculated by MATLAB and Wearable Device

Figure 4. Comparison of features between MATLAB and Wearable Device

In order to further verify the correctness of the feature extraction algorithm of the hardware system, we have conducted an error analysis by comparing features calculated by using MATLAB and that by the wearable device. Table I shows the error analysis of the above four features. In Table I the error range of the energy and entropy is from 0 to 4×10^{-4}, the error range of brightness is from 0 to 10×10^{-3} and the formant error is ranged from 0 to 55Hz. This process verifies the feasibility of speech feature extraction algorithm embedded in the wearable hardware system.

Table I Error Analysis

Audio Features	Absolute Error
Energy	$0 \sim 4\times10^{-4}$
Entropy	$0 \sim 10\times10^{-5}$
Brightness	$0 \sim 10\times10^{-3}$
Formant	$0 \sim 55$

C. *Multi-modal Sensing and Feature Extraction*

Other multi-modal sensing data include 3-axis accelerometer and 3-axis gyroscope sensors (accelerometer measuring range is ±4g and gyroscope measuring range is ±2000°/s) for activity analysis. Atmospheric pressure, ambient temperature and light sensing are used for environment data analysis. Wrist skin temperature and humidity sensing are used for body data analysis. Several features of multi-modal sensing data are calculated, namely mean, standard deviation, short-time energy, energy-entropy, correlation coefficient between axis, pitch, roll and peak value in the frequency domain [5]. In this paper, we have calculated eight features in the time domain and ten features in the frequency domain, such as Mean, STD, Mode, skewness of a FFT window shape, kurtosis of amplitude about a FFT window and so on. Table II describes how these activity behavioral features are calculated.

Table II Feature description for activity analysis

Class	ID	Feature	Description
Time domain	1	Mean	Average value of samples in a window
	2	STD	Standard deviation of samples
	3	Minimum	Minimum of samples in a window
	4	Maximum	Maximum of samples in a window
	5	Mode	The value with the largest frequency
	6	Variance	Variance of samples in a window
	7	Range	Maximum minus minimum
	8	SMA	Signal magnitude area of 3-axis
Frequency domain	9	DC	Direct component of a FFT window
	10-13	Shape Features	Mean, STD, Skewness, Variance of shape about a FFT window
	14-18	Amplitude Features	Mean, STD, Skewness, Variance, Kurtosis of amplitude about FFT

In particular, SMA (Signal Magnitude Area) is the sum of the area enclosed by the triaxial (*x*-, *y*-, *z*-axis) acceleration values, which can make it easy to distinguish between quiet and active states of subjects. The calculation of the SMA can be expressed as:

$$SMA = \frac{1}{t}\left(\int_0^t |x(t)|\,dt + \int_0^t |y(t)|\,dt + \int_0^t |z(t)|\,dt\right) \tag{6}$$

where *t* is the time of samples in a window, $x(t)$, $y(t)$, $z(t)$ is amplitude value of *x*-, *y*-, *z*-axis respectively.

In order to reduce the computational complexity while ensuring accuracy of features, we used synthetic acceleration to calculate the features in Table II. Synthetic acceleration is calculated as:

$$a_i = \sqrt{\left(a_i^x\right)^2 + \left(a_i^y\right)^2 + \left(a_i^z\right)^2} \qquad i \in \{1, 2, \mathrm{L}\ ,\ n\} \tag{7}$$

where a_i is i-th window synthetic acceleration, a_i^x , a_i^y , a_i^z represents triaxial (x-, y-, z-axis) accelerometer readings, respectively.

A sliding window was used for calculating the activity features. The block activity signals are partitioned into short-time periods with a 50%-overlap while each frame consists of 285 sample points. Therefore, the time interval for a set of activity features is 1 second, which is sufficient for capturing behavioral information during for long-time monitoring. As an example, Figure 5 shows four different features in the time and frequency domains, respectively. These features are time mean, FFT kurtosis, SMA and shape std which are obtained from 2.5 hours of accelerometer raw data.

We used the Pearson product-moment correlation coefficient (PPMCC) [51] to calculate the relationship between audio and activity social features with ASQ scores, respectively. The range of correlation value r is from -1 to +1, where 1 represents a total positive linear correlation, 0 represents no correlation and -1 represents a total negative linear correlation. In addition, $0.8 \leqslant |r| \leqslant 1$ represents high correlation, $0.5 \leqslant |r| < 0.8$ represents moderate correlation, $0.3 \leqslant |r| < 0.5$ represents low correlation and $|r| \leqslant 0.3$ represents weak correlation.

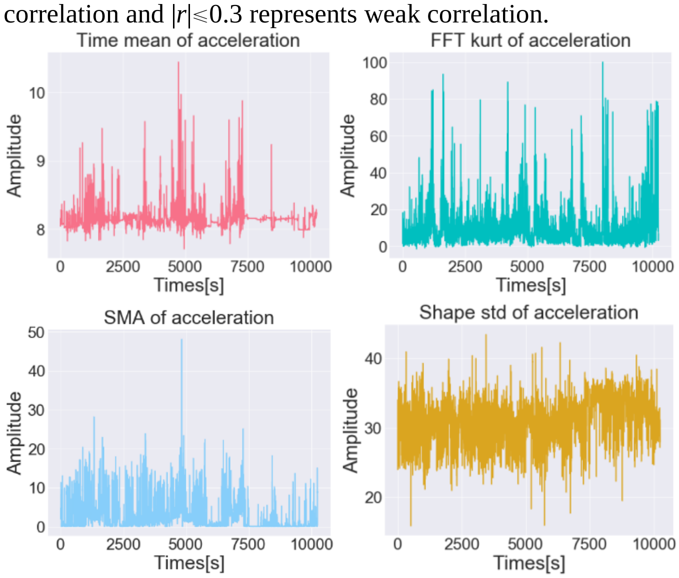


Figure 5. Four different features of time and frequency domain about accelerometer

D. Support Assessment

Subjects were asked to complete a series of questionnaires to assess their personal characteristics including personality and emotional state. Positive and Negative Affect Schedule (PANAS) [52] which is a self-report questionnaire that consists of two 10-item scales to measure both positive and negative affect is used to assess their cognitive and affective well-being. State-Trait Anxiety Inventory (STAI) [53] consists of 40 questions on a self-report basis which can be used to measure two types of anxiety. State anxiety refers how a person is feeling at the time of a perceived threat and is considered as a temporary anxiety. Trait anxiety is a general personal characteristic which measures how much anxiety people feel across typical situations and experiences on a daily basis. Neuroticism Extraversion Openness to Experience Five-Factor Inventory (NEO-FFI) [54] consists of 60 items which can examine a person's personality traits (openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism). The Rosenberg self-esteem scale (RSES) [55] is a self-esteem measurement that is widely used to assess personal characteristics in social-science research. The Beck depression inventory (BDI) [56] is a 21-question multiple-choice self-report inventory and is the most widely used psychometric test for measuring levels of depression. The Autism-Spectrum Quotient (ASQ) [57] for adults (16+ years old) is used for screening autistic symptoms in both healthy subjects and individuals with high-functioning ASD and has a good predictive validity. It is a self-administered questionnaire for measuring the degree to which adults with normal intelligence have traits associated with the autistic spectrum.

E. Experiment

All participants in the study were voluntarily recruited from the University of Electronic Science and Technology of China (UESTC). The volunteers have been required to conduct a pre-mental health related experiments in mental health centers of UESTC. All subjects are right-handed and being free of medical or psychiatric illness, drug or alcohol abuse gave informed consent to take part. In total, 28 students are enrolled from the campus where 16 participants finally joined this experiment (average age is 22.16 ± 0.22 years, 8 undergraduates and 8 Post-graduates; 15 males and 1 female). All the subjects were recruited according to their pre-test ASQ scores. We finally divided all participants into two groups: high-ASQ group (ASQ > 25) and low ASQ group (ASQ < 25).

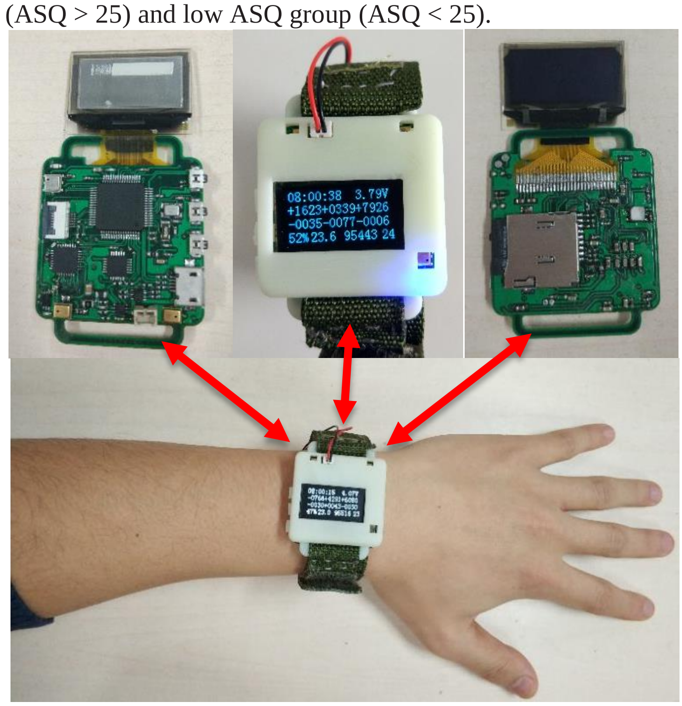
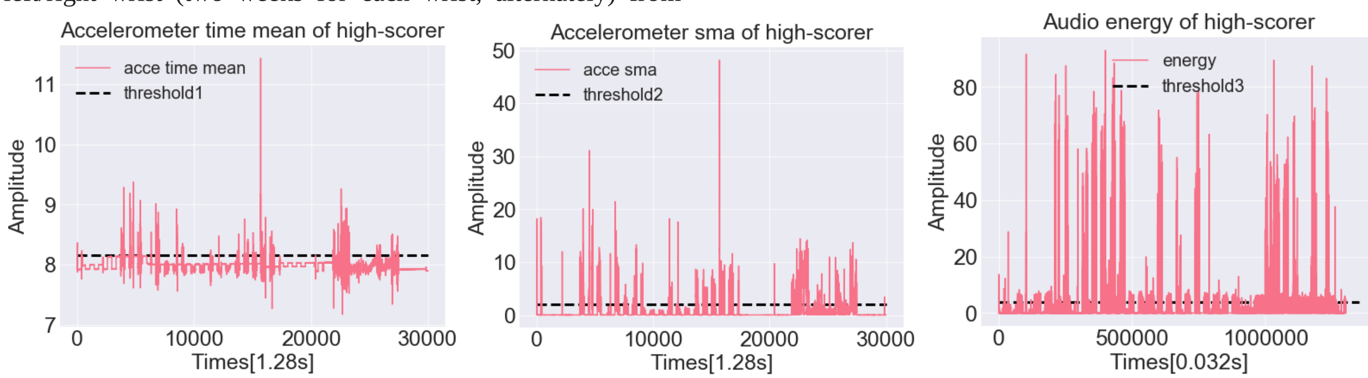


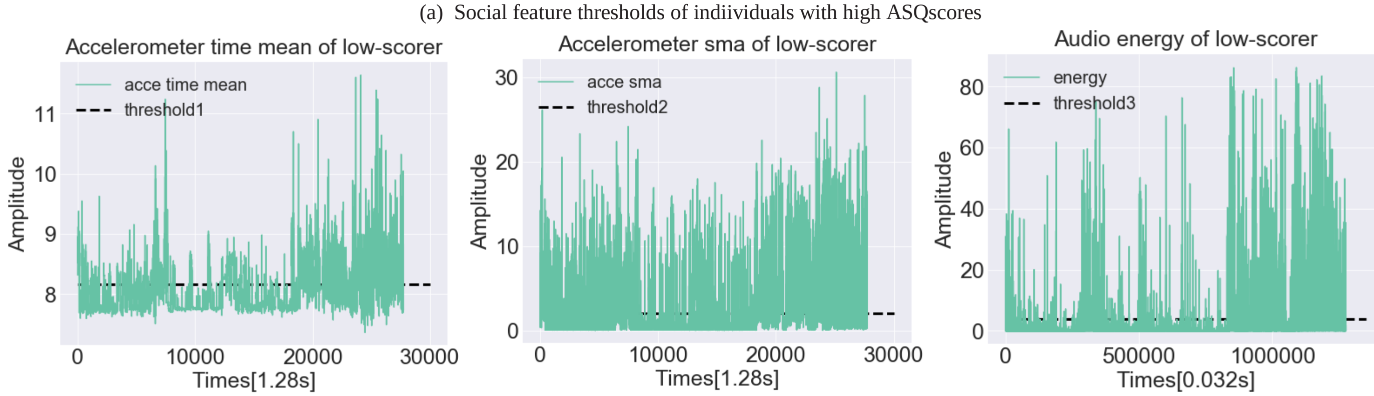
Figure 6. Experimental subject and wearable device

During the first part of the experiment, the subjects were required to come to the laboratory and provide a 10-minute speech including self-introduction and a other language material in a quiet environment and were recorded by the proposed wearable devices as well as an external audio recording device. Then participants were required to complete the PANAS, STAI, NEO-FFI, RSES, BDI and ASQ questionnaires to determine whether they could participate in the experimental study for up to a month. In the 4-week (28 days) experiment studies, participants were required to wear the device constantly and operate it correctly to ensure the integrity of the data. Force majeure or unexpected events in individual cases resulted in temporary suspension of the data collection procedure. In addition, participants were asked to complete an activity log sheet every day. They also needed to repeat a series of the psychological questionnaires including PANAS, SAI, TAI and BDI-II every week in order to explore and changes in their emotional state. Subjects were asked to wear the device on their left/right wrist (two weeks for each wrist, alternately) from morning up to the night for 12 hours without taking it off. Figure 6 shows the internal composition of the wearable device and the way it is worn.

The main objective of the experiment was to use the proposed platform to establish relationships between levels of trait autism in health adults and the long-term daily social sensing features. The ASQ questionnaire scores provided us with an individual's level of autistic symptoms.



(a) Social feature thresholds of indiividuals with high ASQscores



(b) Social feature thresholds of individuals with low ASQ scores

Figure 7. Experimental subject and wearable device

III. Evaluation And Analysis

In this section, we will specify the progress of the social feature analysis.

A. *Threshold selection*

The goal is to extract the information reflecting the social behavior and social communication from the wearable social sensing features. However, long term data collection suffers from strong noise interference and to ameliorate this problem, we calculate the active time through the SMA and energy (accelerometer time mean) of the accelerometer. In addition, the platform estimates the time of speech usage through the short-term energy of audio during a day's monitoring.

Three different thresholds were selected. Specifically, threshold1 is used for SMA, threshold 2 is used for mean time, and threshold 3 is used for short-term energy to reduce the impact of background noise in the long-term monitoring. These thresholds are obtained in a quiet environment. As a typical value, threshold 1 is empirically selected to be 8.16, threshold 2 is empirically selected to be 2 and the threshold 3 is empirically selected to be 4. This facilitates the ease of distinguishing whether the subjects are moving and communicating with others. Figure 7 presents examples where the thresholds are selected to separate the social signals from background noise from ASQ high- and low -scorers. In the figure, it can be clearly deduced that the time of active and speech usage of high-scorer during a day (10 hours) is less (active behavior 51 minutes, speech usage 28 minutes) than that of low-scorer (active behavior 223 minutes, speech usage 52 minutes). The correlation of mental health questionnaires' scores and the social features segmented by these three thresholds during long-term monitoring (1 month, 4 weeks) is significantly improved.

B. *Correlation analysis*

A correlation analysis of 22 social features which are made up of 4 speech features with 13 values including energy, entropy, brightness, 5 formants, the number of formants greater than 0, formants' mean, non-zero formants average, formants' sum, speaking time and 18 activity features with 36 values from 18 accelerometer features and 18 gyroscope features with ASQ scores was performed. In addition, eight statistics were calculated to characterize the social audio and activity features, such as sum, mean, standard deviation, minimum, 25%, 50%, 75% and maximum. Finally, speech usage and ratio of daily activity time using the 3 thresholds was also calculated. In total, 105 audio-related features (i.e.13×8 + 1) and 289 activity-related features (i.e. 36×8 + 1) were calculated. The ASQ scores ranged from 11 to 34.

1). General analysis

In addition to the correlation analysis between social features and ASQ scores, scores from several other mental health and personality questionnaires are analyzed in a similar way by calculating the correlation between audio features, activity features and 12 different psychological questionnaires' scores, respectively.

Table III shows the correlation results between the audio features and the scores of 12 different psychological questionnaires where "numbers" is the number of statistical features. As can be seen, the formant is moderately related to all questionnaires' scores (i.e. formant2_max to ASQ is -0.57, sum_formant_mean to PANAS Positive is -0.65, formant4_mean to PANAS Negative is -0.63, formant_count_25% to SAI is -0.56, sum_formant_mean to TAI is 0.75, formant_mean_min to Neuroticism is 0.61, sum_formant_25% to Extraversion is -0.59, formant_count_75% to Openness is 0.69, sum_formant_mean to Agreeableness is -0.59, formant_mean_min to Conscientiousness is -0.90, formant2_75% to SES is -0.60 and sum_formant_50% to BDI is 0.69). Notwithstanding the above, the social audio features (energy, entropy, brightness and formant) embedded in hardware are associated with at least one questionnaire score (i.e. energy_max to Positive is 0.56, entropy_std to Negative is 0.68, brightness_75% to TAI is -0.62 and formant is correlation with all questionnaires). It can clearly be seen that the number of relevant audio features associated with TAI (19) and Conscientiousness (19) are more than others. The most relevant audio feature for TAI is the sum_formant_mean (0.75) while for Conscientiousness it is the formant_mean_min (-0.90). The reason for this phenomenon may be the different characteristics of ASQ high-scorers and low-scorers for trait anxiety and conscientiousness. Thus, ASQ high-scorers may be more anxious and have a weaker sense of conscientiousness. Indeed, the correlation between ASQ and TAI scores is 0.49. The correlation between TAI and Conscientiousness is -0.78.

Table III Correlation results between audio features and the scores of 12 different psychological questionnaires

Questionnaires	Correlation Features （|Correlation| > 0.5 or |Correlation| > 0.6）		
	|Correlation|	Numbers	Audio Features (correlation value)
ASQ	>0.5	3	sum_formant_50% (0.56), formant2_max (-0.57), formant5_max (-0.53)
Positive	>0.5	6	sum_formant_mean (-0.65), timeTable_75% (0.55), energy_max (0.56) formant2_max (0.56), formant3_max (0.53), formant5_max (0.50)
Negative	>0.5	6	formant4_mean (-0.63), brightness_min (-0.53), sum_formant_max (0.57) formant_count_25% (-0.75), entropy_std (0.68), formant3_75% (0.53)
SAI	>0.5	4	formant1_std (-0.51), brightness_75% (-0.50), formant4_25% (-0.57), formant_count_25% (-0.56)
TAI	>0.6	5	sum_formant_mean (0.75), brightness_75% (-0.62), formant2_min (0.55)
	>0.5	19	entropy_75% (0.52), timeTable_std (-0.59), formant1_max (-0.53)
Neuroticism	>0.5	3	formant_mean_min (0.61), sum_formant_75% (0.51), formant2_75% (0.53)
Extraversion	>0.5	3	energy_std (0.61), timeTable_min (-0.78), sum_formant_25% (-0.59)
Openness	>0.5	6	formant_mean_mean (-0.53), formant_count_75% (0.69), formant1_max (0.52)
Agreeableness	>0.5	3	sum_formant_mean (-0.59), formant2_75% (-0.52), formant_mean_75% (-0.52)
Conscientiousness	>0.6	8	entropy_mean (-0.61), brightness_mean (-0.71), formant2_mean (-0.75)
	>0.5	19	formant_mean5_mean (-0.82), timeTable_std (0.64), formant_mean_min (-0.90)
SES	>0.5	9	formant1_75% (-0.57), formant2_75% (-0.60), formant3_75% (-0.50) formant4_75% (-0.54), energy_std (-0.58), formant_count_50% (0.58)
BDI	>0.5	3	timeTable_75% (-0.53), timeTable_min (0.51), sum_formant_50% (0.69)

Table IV Correlation results between activity features and the scores of 12 different psychological questionnaires

Questionnaires	Correlation Features （|Correlation| > 0.5）	
	Numbers	Activity Features (correlation value)
ASQ	20	acce_time_var_sum (-0.72), acce_fft_mean_sum (-0.69), acce_fft_dc_max (-0.65), acce_shape_var_max (-0.55)
Positive	11	acce_time_std_sum (0.68), acce_fft_mean_sum (0.61), acce_sma_sum (0.63), acce_shape_skew_max (0.51)
Negative	1	ang_time_mode_max (0.62)
SAI	4	acce_time_std_sum (-0.71), acce_time_range_sum (-0.61), acce_fft_mean_sum (-0.66), acce_fft_std_sum (-0.56)
TAI	18	acce_fft_var_mean (-0.65), acce_fft_var_sum (-0.78), timeTable_50% (-0.56), acce_sma_max (-0.59)
Neuroticism	4	acce_fft_var_mean (-0.65), acce_time_mode_std (-0.67), acce_time_var_sum (-0.50), acce_time_mode_max (-0.63)
Extraversion	3	acce_fft_skew_mean (-0.51), acce_fft_kurt_mean (-0.57), ang_time_min_50% (-0.54)
Openness	12	ang_time_var_mean (0.54), ang_time_range_std (0.61), acce_fft_dc_50% (-0.54), ang_fft_mean_max (0.51)
Agreeableness	7	ang_fft_skew_mean (0.51), ang_fft_kurt_mean (0.57), acce_time_mean_25% (-0.56), acce_fft_dc_50% (-0.60)
Conscientiousness	14	acce_fft_std_std (0.67), ang_time_min_std (0.58), ang_time_range_min (0.51), timeTable_50% (0.56)
SES	16	acce_fft_var_mean (0.60), acce_sma_min (0.56), ang_fft_kurt_min (-0.51), ang_time_range_min (0.58)
BDI	4	acce_time_std_sum (-0.53), ang_time_mean_std (0.55), ang_shape_var_min (0.51), acce_fft_var_75% (0.50)

Table IV presents the correlation between the activity features and the scores of 12 different psychological questionnaires. It describes several different activity features where their absolute values of correlation coefficients are greater than 0.5. As can be seen, the time domain and frequency domain features of the accelerometer and gyroscope are sufficient to establish the relationship with mental health and personality questionnaires (i.e. acce_time_var_sum to ASQ is -0.72, acce_shape_skew_max to Positive is 0.51, ang_time_mode_max to Negative is 0.62, acce_fft_mean_sum to SAI is -0.66, acce_fft_var_sum to TAI is -0.78, acce_fft_var_mean to Neuroticism is -0.65, acce_fft_kurt_mean to Extraversion is -0.57, ang_time_range_std to Openness is 0.61, acce_fft_dc_50% to Agreeableness is -0.60, acce_fft_std_std to Conscientiousness is 0.67, ang_fft_kurt_min to SES is 0.51, ang_time_mean_std to BDI is -0.53). The most relevant activity features among the 12 questionnaires is for ASQ in which the number of features above a moderate correlation is 20. This phenomenon indicates that the activity features may well best represent the populations of the different

ASQ scores because ASQ high-scorers are less inclined to move or go out.

2). Speech analysis and ASQ scores

Figure 8 presents the correlation heat map of the audio features and ASQ scores where it clearly indicates that the correlation level of the 6 audio features is around 0.5 (i.e. energy_std is -0.49, sum_formant_50% is 0.57, formant_mean5_50% is 0.57, formant2_max is -0.57, formant3_max is -0.48 and formant5_max is -0.53). The interpretation of these correlation features is listed as follows: energy_std is a standard deviation of the short-term energy, sum_formant_50% is the sum of 5 formants, formant_mean5_50% is a mean of the 5 formants, formant2_max is the maximum of the 2nd formant. The obtained results clearly indicate that social activity features of energy and formant are able to distinguish between the ASQ high- and low-scoring groups.

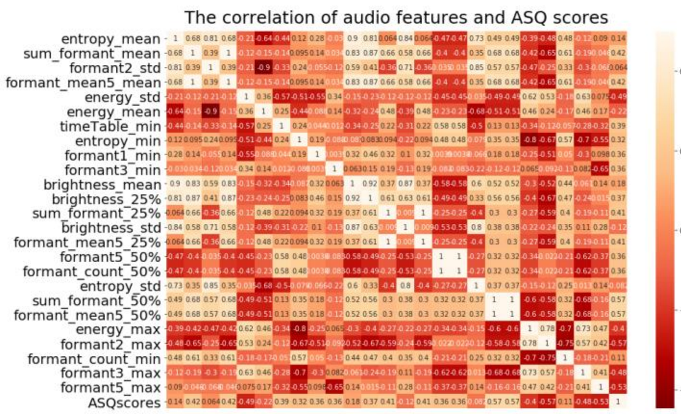


Figure 8. Correlation heat map of audio features and ASQ scores

3). Activity analysis and ASQ scores

Figure 8 presents the correlation heat map of the audio features and ASQ scores. Both time and frequency domain features are strongly correlated with ASQ (i.e. acce_time_var_std is -0.73 and acce_fft_mean_sum is -0.69). The reason for the high correlation between behavioral features and ASQ scores has been hypothesized to be due to their different living habits where ASQ high-scorers tend to prefer sitting and nonsocial activities while low-scorers are more willing to participate in social activity and communication with others. The most related time domain features of accelerometer for ASQ are listed as follows: acce_time_mean_max is -0.65, acce_time_std_sum is -0.59, acce_time_max_max is -0.63, acce_time_mode_std is -0.61, acce_time_var_std is -0.73, acce_time_range_sum is -0.71, acce_sma_max is -0.65.

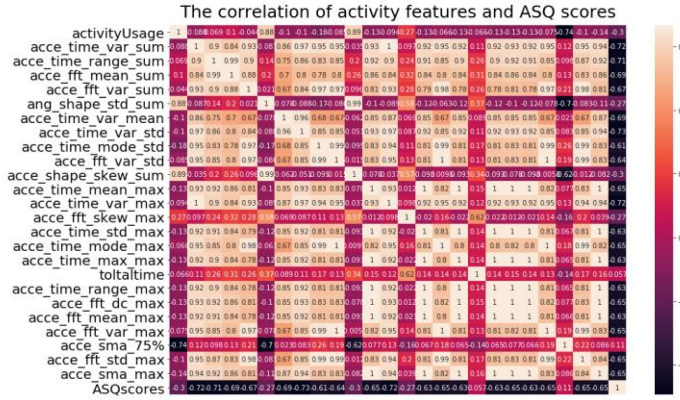
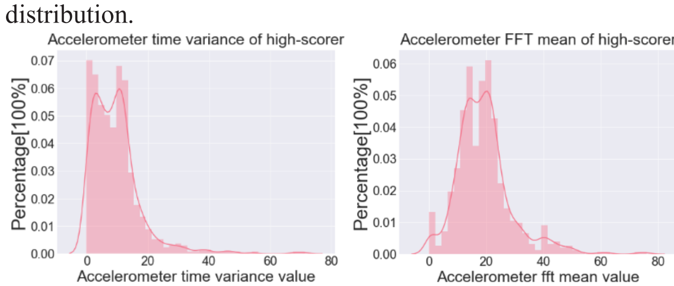


Figure 9. Correlation heat map of activity features and ASQ scores

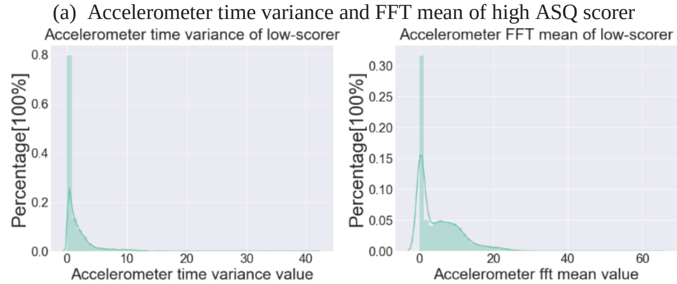
C. *Long time analysis*

1). High ASQ group vs low ASQ group in a day

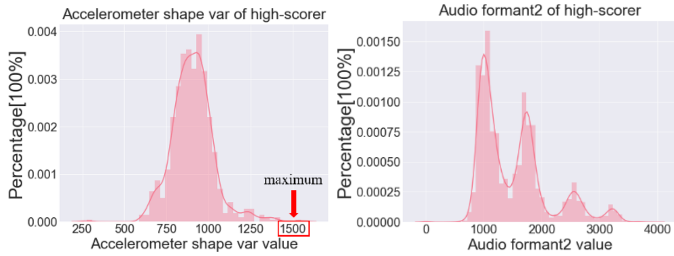
In order to compare and analyze the differences in social features between the high ASQ group and the low ASQ group, the Kernel Density Estimation (KDE) is calculated and the density curves of the social features have been drawn. These are accelerometer time variance, accelerometer FFT mean, accelerometer shape variance and the 2nd audio formant during a day as shown in Figure 10. It can be clearly seen that the 80% acce_time_var value of the subjects from the low ASQ group is close to 0 while the value of the counterparts from the high ASQ group is approximately gaussian distributed between 0 and 40. For accelerometer FFT mean (frequency domain feature) between subjects in the two groups, there is a similar trend to accelerometer time variance (time domain feature). Although the density curses of accelerometer shape variance between the subjects in the two groups are almost identical, the maximum value of high-scorers is approximate 1500 while the maximum value of low-scorers is approximate 2000 (see (c) and (d)). In addition, there are significant differences in the 2nd audio formant between the subjects in two groups during the day. The density curve of the subjects in the high ASQ group tends to a sparse distribution from 900Hz to 3500Hz while the density curve for the subjects in low ASQ group tends to a Gaussian distribution.



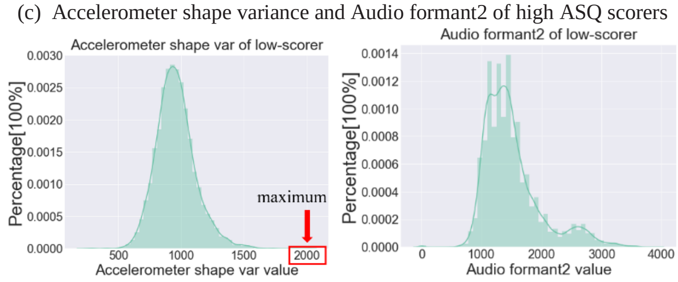
(a) Accelerometer time variance and FFT mean of high ASQ scorer



(b) Accelerometer time variance and FFT mean of low ASQ scorer



(c) Accelerometer shape variance and Audio formant2 of high ASQ scorers



(d) Accelerometer shape variance and Audio formant2 of low ASQ scorer

Figure 10. The density curves of social features during a day

2s). ASQ high-scorers vs ASQ low-scorers during a month

Figure 11 shows the tendency curves of the audio *2*nd formant max and accelerometer time mean max of ASQ high-scorers and low-scorers during a month. As can be seen, the frequency range of the low-scorer's formant2_max is from 3650 Hz to 3950Hz while that for the high-scorer's the frequency range is from 3200Hz to 3950Hz which fluctuates with more volatility. In addition, the acce_time_mean_max has a similar tendency with the low-scorer's amplitude varying dramatically while the high-scorer is more constant. Overall, the audio *2*nd formant max and accelerometer time mean max average of the low-scorer are higher than that of the high-scorer.

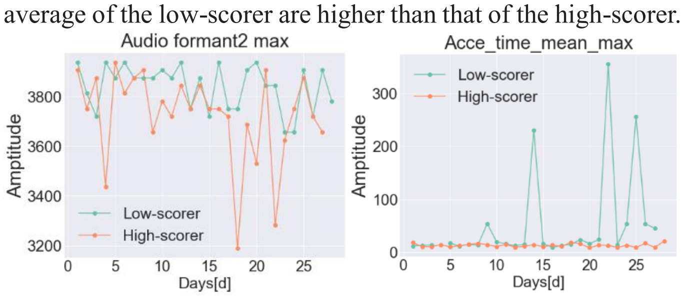


Figure 11. Tendency curves of social features during a month

IV. CONCLUSION

In this paper, a portable wearable device has been designed and implemented for long-term person wellbeing monitoring. In particular, four audio-wellbeing features have been embedded into a wearable hardware platform without preserving the raw audio data to avoid privacy issues. The feasibility of the study has been verified through a four-week long-term monitoring experimental study and the correlations of the social features and the questionnaires scores have been calculated. The results obtained confirmed that at least one social audio feature (formant) is highly correlated with questionnaire scores. In addition, there are significant differences in the accelerometer time domain features and the *2*nd audio formant between individuals with high as opposed to low autism traits (ASQ scores).

REFERENCES

[1] Picard, Rosalind W. "Affective computing: challenges." *International Journal of Human - Computer Studies* 59.1(2003):55-64.

[2] Vinciarelli, Alessandro, M. Pantic, and H. Bourlard. "Social signal processing: Survey of an emerging domain." *Image & Vision Computing*27.12(2009):1743-1759.

[3] Thoits, P. A. "Mechanisms linking social ties and support to physical and mental health. " *J Health Soc Behav* 52.2(2011):145-161.

[4] Bin Gao, and Wai Lok Woo. "Wearable Audio Monitoring: Content-Based Processing Methodology and Implementation." *IEEE Transactions on Human-Machine Systems* 44.2(2014):222-233.

[5] Gu, Jun, et al. "Wearable Social Sensing: Content-Based Processing Methodology and Implementation." *IEEE Sensors Journal* PP.99(2017):1-1.

[6] Pantelopoulos, Alexandros, and N. G. Bourbakis. "A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis." *IEEE Transactions on Systems Man & Cybernetics Part C* 40.1(2009):1-12.

[7] Yatani, Koji, and K. N. Truong. "BodyScope:a wearable acoustic sensor for activity recognition." *ACM Conference on Ubiquitous Computing* ACM, 2012:341-350.

[8] Pantelopoulos, Alexandros, and N. G. Bourbakis. "A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis." *IEEE Transactions on Systems Man & Cybernetics Part C* 40.1(2009):1-12.

[9] Albrecht, S. "Advanced interaction in context." *Proc.int.symp.on Handheld & Ubiquitous Computing* 1707.5(1999):89-101.

[10] Farringdon, J. "Wearable Sensor Badge & Sensor Jacket for Context Awareness." *Proc Iswc* (1999).

[11] Choudhury, Tanzeem, et al. "The Mobile Sensing Platform: An Embedded Activity Recognition System." *IEEE Pervasive Computing* 7.2(2008):32-41.

[12] Olguin, Olguin D, et al. "Sensible organizations: technology and methodology for automatically measuring organizational behavior. " *IEEE Transactions on Systems Man & Cybernetics Part B Cybernetics*39.1(2009):43-55.

[13] Nag, Anindya, S. C. Mukhopadhyay, and J. Kosel. "Wearable Flexible Sensors: A Review." *IEEE Sensors Journal* 17.13(2017):3949-3960.

[14] Patel, Shyamal, et al. "A review of wearable sensors and systems with application in rehabilitation." *J Neuroeng Rehabil.* 9.1(2012):21

[15] Teng, Xiao Fei, et al. "Wearable Medical Systems for p-Health." *IEEE Rev Biomed Eng* 1(2008):62-74.

[16] Bonato, P. "Wearable sensors and systems. From enabling technology to clinical applications. " *IEEE Engineering in Medicine & Biology Magazine the Quarterly Magazine of the Engineering in Medicine & Biology Society*29.3(2010):25.

[17] Torrado, Juan C., J. Gomez, and G. Montoro. "Emotional Self-Regulation of Individuals with Autism Spectrum Disorders: Smartwatches for Monitoring and Interaction." *Sensors* 17.6(2017):1359.

[18] Fletcher, Richard R., M. Z. Poh, and H. Eydgahi. "Wearable sensors: Opportunities and challenges for low-cost health care." *Engineering in Medicine and Biology Society* IEEE, 2010:1763-6.

[19] Bekele, E., et al. "Multimodal adaptive social interaction in virtual environment (MASI-VR) for children with Autism spectrum disorders (ASD)." *Virtual Reality* IEEE, 2016:121-130.

[20] Ouchi, K, T. Suzuki, and M. Doi. "LifeMinder: a wearable healthcare support system using user's context." *International Conference on Distributed Computing Systems* IEEE Computer Society, 2002:791-792.

[21] Hong, Yu Jin, et al. "Activity Recognition Using Wearable Sensors for Elder Care." *International Conference on Future Generation Communication and NETWORKING* IEEE, 2008:302-305.

[22] Jiang, Ping, et al. "An Intelligent Information Forwarder for Healthcare Big Data Systems With Distributed Wearable Sensors." *IEEE Systems Journal* 10.3(2016):1147-1159.

[23] Patel, Shyamal, et al. "A review of wearable sensors and systems with application in rehabilitation." *J Neuroeng Rehabil.* 9.1(2012):21.

[24] Fletcher, R. R., et al. "Wearable sensor platform and mobile application for use in cognitive behavioral therapy for drug addiction and PTSD. " *Engineering in Medicine and Biology Society, Embc, 2011 International Conference of the IEEE* IEEE, 2011:1802.

[25] Howell, Jonathan, et al. "A low-power wearable substance monitoring device." *Applications of Commercial Sensors* IEEE, 2016:1-9.

[26] Wang, J. B., et al. "Wearable Sensor/Device (Fitbit One) and SMS Text-Messaging Prompts to Increase Physical Activity in Overweight and Obese Adults: A Randomized Controlled Trial." *Telemedicine journal and e-health: the official journal of the American Telemedicine Association*21.10(2015):782-92.

[27] Kim, Jayoung, et al. "Non-invasive Alcohol Monitoring Using a Wearable Tattoo-based Iontophoretic-Biosensing System." (2016).
[28] Norris, R, D. Carroll, and R. Cochrane. "The effects of physical activity and exercise training on psychological stress and well-being in an adolescent population." *Journal of Psychosomatic Research*36.1(1992):55-65.
[29] Lin, Mu, et al. "BeWell+: multi-dimensional wellbeing monitoring with community-guided user feedback and energy optimization." *Conference on Wireless Health* ACM, 2012:1-8.
[30] Wilhelm, F. H., and W. T. Roth. "The somatic symptom paradox in DSM-IV anxiety disorders: suggestions for a clinical focus in psychophysiology. " *Biological Psychology* 57.1-3(2001):105.
[31] Roh, T, et al. "Wearable mental-health monitoring platform with independent component analysis and nonlinear chaotic analysis." *Engineering in Medicine and Biology Society* IEEE, 2012:4541-4544.
[32] Vidal, Mélodie, et al. "Wearable eye tracking for mental health monitoring." *Computer Communications* 35.11(2012):1306-1311.
[33] Lanata, A, et al. "Complexity index from a personalized wearable monitoring system for assessing remission in mental health." *IEEE J Biomed Health Inform* 19.1(2015):132-139.
[34] Lu, Hong, et al. "StressSense: detecting stress in unconstrained acoustic environments using smartphones." *Ubicomp '12 Acm* (2012):351-360.
[35] Naslund, J. A., K. A. Aschbrenner, and S. J. Bartels. "Wearable Devices and Smartphones for Activity Tracking Among People with Serious Mental Illness." *Mental Health & Physical Activity* 10(2016):10.
[36] Wilhelm, Frank H., and P. Grossman. "Emotions beyond the laboratory: Theoretical fundaments, study design, and analytic strategies for advanced ambulatory assessment." *Biological Psychology*84.3(2010):552-569.
[37] Mast, M. Schmid, et al. "Social Sensing for Psychology: Automated Interpersonal Behavior Assessment." *Current Directions in Psychological Science* 24.2(2015):154-160.
[38] Kakria, Priyanka, N. K. Tripathi, and P. Kitipawang. *A Real-Time Health Monitoring System for Remote Cardiac Patients Using Smartphone and Wearable Sensors*. Hindawi Publishing Corp. 2015.
[39] Gaggioli, A, and G. Riva. "From mobile mental health to mobile wellbeing: opportunities and challenges. " *Studies in Health Technology & Informatics* 184(2013):141.
[40] Farchi, S, et al. "Activity-aware Mental Stress Detection Using Physiological Sensors." 23.13(2010):211-230.
[41] Ertin, Emre, et al. "AutoSense: unobtrusively wearable sensor suite for inferring the onset, causality, and consequences of stress in the field." *ACM Conference on Embedded Networked Sensor Systems* ACM, 2011:274-287.
[42] Tseng, Vincent W. S., et al. "Assessing mental health issues on college campuses: preliminary findings from a pilot study." *ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* ACM, 2016:1200-1208.
[43] Sano, A, et al. "Recognizing Academic Performance, Sleep Quality, Stress Level, and Mental Health using Personality Traits, Wearable Sensors and Mobile Phones." *IEEE, International Conference on Wearable and Implantable Body Sensor Networks* IEEE, 2015:1-6.
[44] Leonard, N. R., et al. "Mobile Health Technology Using a Wearable Sensorband for Female College Students With Problem Drinking: An Acceptability and Feasibility Study:." *Jmir Mhealth Uhealth* 5.7(2017):e90.
[45] Jaques, Natasha, et al. "Predicting students' happiness from physiology, phone, mobility, and behavioral data." *International Conference on Affective Computing and Intelligent Interaction* IEEE Computer Society, 2015:222-228.
[46] Wang, Rui, et al. "StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones." *ACM International Joint Conference on Pervasive and Ubiquitous Computing* ACM, 2014:3-14.
[47] Perez, Alfredo, and S. Zeadally. "Privacy Issues and Solutions for Consumer Wearables." *It Professional* PP.99(2017):1-1.
[48] Nass, C., and S. Brave. "Voice Activated: How People Are Wired for Speech and How Computers Will Speak with Us." (2004).
[49] Pentland, Alex. "Social Dynamics: Signals and Behavior." *International Conference on Developmental Learning* 2004:263--267.
[50] Ambady, Nalini, and R. Rosenthal. "Thin Slices of Expressive Behavior as Predictors of Interpersonal Consequences: A Meta-Analysis." *Psychological Bulletin* 111.2(1992):256-274.
[51] Hauke, Jan, and T. Kossowski. "Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data." *Quaestiones Geographicae* 30.2(2011):87-93.
[52] Watson, David, L. A. Clark, and A. Tellegen. "Development and validation of brief measures of positive and negative affect: The PANAS scales. " *J Pers Soc Psychol* 54.6(1988):1063-70.
[53] Spielberger, Cd. "STAI manual for the State-trait anxiety inventory." *Self-Evaluation Questionnaire* iv(1970):1-24.
[54] Mccrae, R. R. "The NEO-PI/NEO-FFI manual supplement." (1989).
[55] Rosenberg, Morris. "Self Esteem and the Adolescent. (Economics and the Social Sciences: Society and the Adolescent Self-Image)." *New England Quarterly* 148.2(1965):177-196.
[56] Whisman, M. A., J. E. Perez, and W. Ramel. "Factor structure of the Beck Depression Inventory—Second Edition (BDI-ii) in a student sample." *Journal of Clinical Psychology* 56.4(2000):545-551.
[57] Baron-Cohen, S., et al. "The autism-spectrum quotient (AQ): Evidence from Asperger Syndrome/high-functioning autism, males and females, scientists and mathematicians (vol 31, pg 5, 2001)." *Journal of Autism & Developmental Disorders* 31.6(2001):603-603.