Proceedings Paper:

Flanagan, D., Arvaneh, M. and Zaffaroni, A. (2014) Audio signal analysis in combination with noncontact bio-motion data to successfully monitor snoring. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014. EMBC , 26-30 Aug 2014, Chicago, USA. IEEE , pp. 3763-3766.

https://doi.org/10.1109/EMBC.2014.6944442

Audio Signal Analysis in Combination with Noncontact Bio-motion Data to Successfully Monitor Snoring

David Flanagan, Mahnaz Arvaneh, Alberto Zaffaroni

***Abstract*-- This paper proposes a novel algorithm for automatic detection of snoring in sleep by combining non-contact bio-motion data with audio data. The audio data is captured using low end Android Smartphones in a non-clinical environment to mimic a possible user-friendly commercial product for sleep audio monitoring. However snore detection becomes a more challenging problem as the recorded signal has lower quality compared to those recorded in clinical environment. To have an accurate classification of snore/non-snore, we first compare a range of commonly used features extracted from the audio signal to find the best subject-independent features. Thereafter, bio-motion data is used to further improve the classification accuracy by identifying episodes which contain high amounts of body movements. High body movement indicates that the subject is turning, coughing or leaving the bed; during these instances snoring does not occur. The proposed algorithm is evaluated using the data recorded over 25 sessions from 7 healthy subjects who are suspected to be regular snorers. Our experimental results showed that the best subject-independent features for snore/non-snore classification are the energy of frequency band 3150-3650 Hz, zero crossing rate and 1st predictor coefficient of linear predictive coding. The proposed features yielded an average classification accuracy of 84.35%. The introduction of bio-motion data significantly improved the results by an average of 5.87% ($p<0.01$). This work is the first study that successfully used bio-motion data to improve the accuracy of snore/non-snore classification.**

I. Introduction

Snoring is more than just a nocturnal nuisance. Indeed, it can be regarded as a major sign of sleep disorder prevalence. Snoring is a consistent symptom leading to one of the most common sleep disorders, known as sleep apnoea. Snoring is prevalent in a substantial percentage of males and females [1]. The commonly accepted percentages of snorers are 40% in males and 20% in females [2].

Automatic detection of snoring using audio signals can characterize a subject's snores. This information can benefit the subject by alerting them if they are snoring consistently and recommend them to contact and seek help from a clinician about a possible sleep disorder they may suffer from. To automate the detection of snoring there are several challenges to overcome, the primary challenge will be to differentiate snoring audio from other noise such as heavy breathing, body turning, sleep talking noise, external noise etc. During the data collection the position of the recording device may change from session to session. This will result in variation in properties of the audio signal causing difficulty in detecting snore from session to session.

There are several studies attempting to automatically detect snores. One study used energy and Zero Crossing Rate (ZCR) to identify snoring frames from audio signals recorded using a microphone hung above a subject from the ceiling in a clinical lab [3]. Another study used the Hidden Markov Model and spectral based features for automatic segmentation and classification of different types of sounds such as snoring, breathing and others [5]. Yadollahi et al. used zero-crossing rate, logarithm of the signal's energy and first format frequency as characteristic features in [4]. They used two microphones for the data collection, one placed on the subject's throat to record the tracheal and one ambient microphone was used. Another study used a combination of zero crossings, energy of signal, normalized autocorrelation coefficient and the first predictor coefficient of Linear Predictive Coding (LPC) analysis to differentiate silence, snoring and other classes of sound. The recording method used a microphone which was hung from the ceiling above the subject [6].

Despite various studies proposing different features for snore detection, the best subject-independent features which work reliably across all subjects have not yet been investigated. Moreover, most of the past studies used high range dictaphones and recording equipment to record the audio. They conducted their data collection in clinical environments with the recording microphones in a constant position from session to session [3-9]

Unlike the past studies, this paper uses a low end smartphone to collect the data in the user's home environment. The difference in data collection mimics a possible user-friendly commercial product environment, although it also introduces much more artifacts, noise and non-stationaries making the snore detection more challenging. To have an accurate classification of snore/non-snore, this paper, first, compares a range of previously proposed snore detection features, and finds the best features working reliably across all the subjects. Thereafter, data from a non-contact bio-motion sensor is used to further improve the classification accuracy. To the best of our knowledge, this is the first study using non-contact bio-motion sensor data to improve classification of snore/non-snore detection algorithms.

The experimental protocol records audio data using an Android Smartphone, while a non-contact bio-motion sensor records body movement data. The non-contact bio-motion sensor senses the movement of a subject using an ultra-low-power radio-frequency transceiver that sends and receives

D. Flanagan is with School of Electrical, Electronic and Communications Engineering, University College Dublin, Dublin 4, Ireland (email: david.flanagan.1@ucdconnect.ie).

M. Arvaneh is with University College Dublin and Trinity College Neuroscience Institute, Dublin, Ireland (e-mail: mahnaz.arvaneh@ucd.ie).

A. Zaffaroni is with ResMed Sensor Technologies, NexusUCD, Blocks 9 & 10, Belfield Office Park, University College Dublin, Belfield, Dublin 4. Ireland (e-mail: alberto.zaffaroni@resmed.com).

radio waves. The proposed algorithm is evaluated using data collected from 7 snoring subjects which consist of a total 25 sleep sessions.

II. Experiment

In this study, the experiment participants, guided with an information leaflet, conducted the data collection trial in their own home environment with the position of the recording smartphone at their own discretion. Thus, the sleep signals were recorded in a real life situation mimicking how possible future commercial users would use a smartphone to record themselves sleep. However, this introduced a variety of external noise and artifacts which led to inter and intra-session variability. Thus compared to previous studies, the proposed snore detection algorithm was required to handle these new challenges.

A. Participants

The experiment participants were all healthy subjects with no history of sleep disorders. The participants had an observational history of snoring. Seven subjects took part in the experiment; this resulted in 25 sleep audio and bio-motion sleep recordings. Three of the seven subjects had sleeping partners.

B. Audio Data Collection:

A participant is given an Android Smartphone, *Alcatel V860* [10], with a recording application installed on it (i.e. *Sound & Voice Recorder – ASR* [11]). The Android application records in mono in .WAV format with a sampling rate of 8000 Hz, and a bit rate of 64 kbps. The participants are required to open the recording application each night and initiate recording, followed by placing the smartphone on their bedside locker or mattress. The exact position is at the participant's discretion. After finishing the sleep, the participant opens the recoding application, and stops the application recording. The audio file for the previous night's sleep is saved on the external SD card of the smartphone.

C. Non-contact Bio-motion Sensor Data Collection

The non-contact bio-motion sensor, used in this study, is a device designed and developed by ResMed Sensor Technologies. The system is based on a novel non-contact bio-motion sensor which uses an ultra-low-power radio-frequency transceiver to send and receives radio waves to sense the movement of a subject [12]. A major advantage of an RF signal is that it can work through blankets, bedding sheets etc. The device logs the data at a rate of 64 samples/second to an external SD card. The data is logged with 12 bits ADC precision. The greater the ac shift in the signal, the larger the subject movement.

A trial participant is given a non-contact bio-motion sensor. Initially, the participant is instructed to place the device on their bedside table with the front of the device having a clear line of sight to the participant's chest. The participant leaves the bio-motion sensor powered on for the duration of the trial. Introducing the body movement reading from the non-contact bio-motion sensor to improve the accuracy of a snore algorithm is a novel idea.

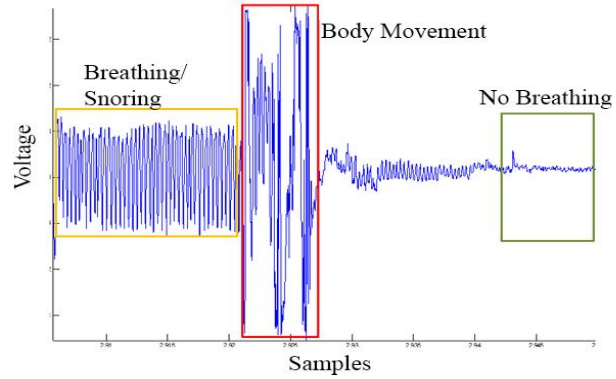

Figure 1. Example of the bio-motion sensor data

III. Proposed Snore Detection Algorithm

A. Sound Episode Extraction

The audio signals are windowed to a length of 120ms, with 50% overlap as suggested by [5]. The first step in the proposed snore detection algorithm is to eliminate silence and segment intervals of sound activity. The energy of a full recording is calculated. A simple thresholding criterion as proposed in [13] is used to eliminate silence areas of the audio signal. Energy levels below the threshold are set to 0. This method creates episodes of sound activity separated by 0 energy data points corresponding to the silence.

B. Feature Selection

The next step of the proposed algorithm is to find the best features across all the subjects. Using training data extracted from the recorded data, 13 different features, commonly used in literature, were investigated [3-7, 10]. The investigated features are as follows:

1-8*)* Normalized averaged energy of each window extracted using a 500 Hz sub-band (obtained by dividing the 150–4150 Hz frequency range into eight 500 Hz sub-bands)
9) Duration of each sound episode
10) Maximum ZCR of all windows in each sound episode
11) Sum of energy of each sound episode
12) 1^{st} predictor coefficient of LPC of each sound episode
13) Mean of Spectral Centroids of all windows in each sound episode

For the first 8 features, normalization was done by dividing the energy of the window extracted from the investigated sub-band to the total energy of the window.

Fisher scores [14], correlation coefficients and t-test ranking (i.e. ranking each feature based on the p-value between the samples of the two classes) algorithms were used to rank the features. The algorithms were applied subject-independently to rank the entire pool of the subject training data based on their class labels. Based on each algorithm, the features were ranked. The five highest ranked features obtained by averaging the results of all the three abovementioned ranking algorithms are as follows:

1. 1st predictor coefficient of LPC of each sound episode
2. Normalized averaged energy obtained from the 7th frequency sub-band (3150-3650 Hz)
3. Normalized averaged energy obtained from the 6th frequency sub-band (2650-3150 Hz)
4. Normalized averaged energy obtained from the 8th frequency sub-band (3650-4150 Hz)
5. Maximum ZCR of all windows in each sound episode

The results obtained by the three ranking algorithms are presented in detail in Section IV.

Since the proposed ranking algorithms rank the features individually without considering the correlation between them, we applied the 10 fold-cross validation method on the train data to find the best combination among the 5 highest ranked features. Several classifiers were tested using the 5 best features found. Different combinations and permutations were analyzed. The bio-motion sensor data was not used in this part. We limited ourselves to only check the combinations of the 5 top ranked features rather than all the 12 features to save on computation time.

The K-Nearest Neighbor (KNN) classifier with $K = 3$ resulted in the highest overall accuracy using three of the 5 best features. The results, obtained using the KNN classifier, are presented in Section IV. The 3 optimum features are as follows:

1. 1st predictor coefficient of LPC of each sound episode
2. Normalized averaged energy obtained from the 7th frequency sub-band (3150-3650 Hz)
3. Maximum ZCR of all windows in each sound episode

C. *Bio-motion Sensor Data*

The bio-motion sensor data and audio data are recorded on two separate devices. Initially, the audio and bio-motion data is synchronised using the corresponding device timestamps with an accuracy of ±1 second of each other. As the bio-motion sensor device has a sampling rate of 64Hz, the data is up sampled to coincide with the audio signal's sampling rate of 8000Hz.

Two methods were analysed to manipulate the bio-motion sensor data to improve the classification phase of the proposed algorithm:

1. Bio-motion data as a classification input feature (abbreviated as BM_FT)
2. Binarized bio-motion data based on a threshold as a classification input feature (abbreviated as BM_TH_FT)

The bio-motion signal is centred on zero and is windowed using the same window parameters as the audio data. The maximum magnitude of the signal in each window is then calculated. This magnitude is used as a classification input feature in the BM-FT method.

In the BM_TH_FT method, after centring the bio-motion data, a bio-motion threshold value is taken as the 99 percentile value of the absolute bio-motion signal. The maximum magnitude of the signal in each window is compared to the threshold. If the value is below the threshold the bio-motion feature for that episode is set to zero, otherwise the value is 1. Fig. 2 illustrates the proposed snore/non-snore detection algorithm that uses BM_TH_FT as a feature.

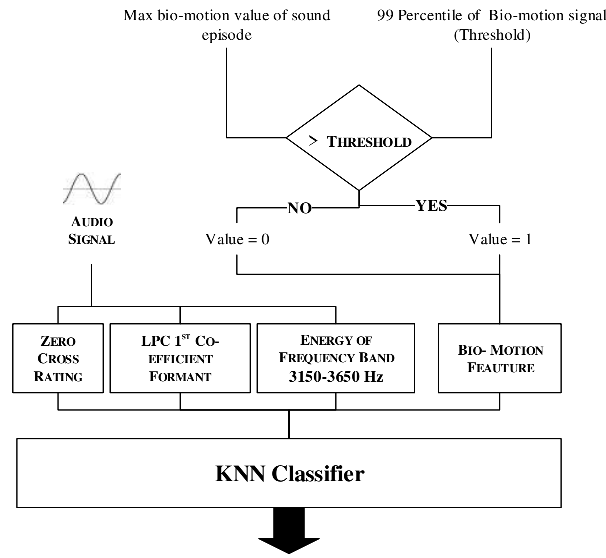

Figure 2. A diagram presenting our proposed snore detection algorithm using the BM_TH_FT method

IV. EXPERIMENTAL RESULTS

A. *Training and Test Data*

For classification and feature analysis, sections of the sleep recordings were listened to and training data was manually annotated from each sleep recording. To test the classifiers, test data was created by taking 20% of the training data from each session.

B. *Best Feature Selection*

Three ranking algorithms where used to rank the best features. The Fisher Score algorithm (FS) was implemented using weighted least squares regression routines [14]. The T-test rank (T-T) is based on a pool variance estimate. Lastly, the correlation co-efficient between the training data and the class labels was calculated to rank the features (CC).

Table I presents the results of the three ranking algorithms. The features have been ranked from 1-13 with 1 indicating the best ranked feature for that method. Table 1 shows that feature number 12 (i.e. 1st predictor coefficient of LPC of each sound episode) is the best feature based on the CC and FC ranking algorithms. On the other side, T-T algorithm identified the feature number 7 (i.e. normalized averaged energy obtained from 3150-3650 Hz) as the best feature. Averaging across the results of the three ranking algorithms showed that the overall top five ranked features among the 13 features described in Section III.B are the features 12, 7, 8, 6 and 10.

To differentiate between snore and non-snore sound episodes, several methods of classification were analysed using varied combinations of the selected features. The results indicate that the most accurate classifier for our problem is a KKN classifier using the features ZCR, 1st predictor coefficient of LPC, and energy of frequency sub-band 3150-3650 Hz.

TABLE I. FEATURE RANKING USING THE THREE RANKING ALGORITHMS. THE FEATURES WERE INDICATED BASED ON A NUMBER DESCRIBED IN SECTION III.B.

	Corresponding Features Numbers (Defined In Section III.B)												
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Method	Ranking Values of Corresponding Features (1=Best Feature)												
CC	7	12	10	13	11	5	6	8	2	9	14	1	3
FS	13	9	10	7	6	5	4	3	11	8	2	1	14
T-T	13	11	14	12	4	5	1	3	9	2	6	8	7
Avg.	11	10.7	11.3	10.7	7	5	3.7	4.7	7.3	6.3	7.3	3.3	8
Overall	13	11	14	11	6	**4**	**2**	**3**	7	**5**	7	**1**	9

C. *Effect of Bio-motion Data on Classification Accuracy*

To consider the effect of bio-motion data on improving the classification accuracy, three different models were trained using the train data of each subject and evaluated using the test data of the same subject. The first model only used the best three acoustic features, introduced in Section III.B, for training the KNN classifier. This model is abbreviated as (W/O_BM). The two remaining models used one of the defined bio-motion features in addition to the best three acoustic features. These models were abbreviated as BM_FT (i.e. Bio-motion data used directly as a feature) and BM_TH_FT (i.e. Binarized bio-motion data based on a threshold used as a feature).

As shown in Table II and Fig. 3 the use of the bio-motion data as either BM-FT or BM-TH-FT significantly improved the snore/non-snore detection algorithm. The proposed BM-TH-FT algorithm yielded an average of 1.9% improvement compared to W/O-BM. Interestingly, the proposed BM-FT algorithm outperformed the W/O-BM algorithm by an average of 5.87%. The paired t-test showed that this improvement is statistically significant ($p<0.01$).

TABLE II. CLASSIFICATION EFFECT OF BIO-MOTION FEATURES

Method	**Avg. Accuracy ± std**	**Avg. Classification effect (%)**	**T-Test Score****
W/O_BM	84.3 ± 9.2	0	-
BM_FT	90.3 ± 8.1	+5.87	0.00001
BM_TH_FT	86.3 ± 9.7	+1.93	0.007

** T-Test score obtained in comparison with W/O_BM Results

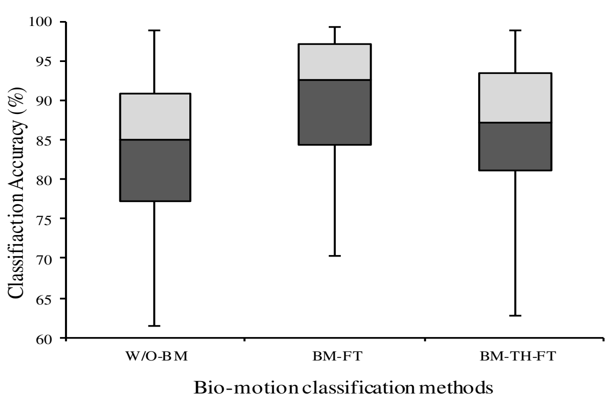

Figure 2. Snore detection results using three proposed algorithms

V. CONCLUSION

This paper indicated that recorded audio signals from a smartphone in combination with data from a non-contact bio-motion sensor could be translated into a user-friendly commercial product to screen a subject's snore. For this purpose, first 13 different combinations of commonly used acoustic features in snore detection were investigated and the three best features that work reliably across all the subjects were selected. Importantly, introducing the bio-motion data as a new feature to the classifier yielded an average improvement of 5.87% which was statistically significant ($p<0.01$). To further expand this study it would be interesting to investigate a subject-independent snore detection classifier designed based on acoustic and bio-motion data.

VI. REFERENCES

[1] E. Lugaresi, F. Cirignotta, G. Coccagna and C. Piana, "Some epidemiological data on snoring and cardiocirculatory disturbances," *Sleep.,* vol. 3, no. 3, pp. 221-4, 1980.

[2] H. Bearpark, L. Elliott, R. Grunstein, S. Cullen, H. Schneider, W. Althaus and C. Sullivan, "Snoring and sleep apnea. a population study in australian men," *American Journal of Respiratory and Critical Care Medicine,* vol. 151, no. 5, pp. 1459-1465, 1995.

[3] U. Abeyratne, A. Wakwella and C. Hukins, "Pich jump probability measures for the analysis of snoring sounds in apnea," *Physiol Meas,* vol. 26, no. 5, pp. 779-798, 2005.

[4] A. Yadollahi and Z. Moussavi, "Automatic breath and snore sound classification from tracheal and ambient sounds recordings," *Medical Engineering and Physics,* vol. 32, no. 9, pp. 985-90, 2010.

[5] W. Duckitt, S. Tuomi and T. Nielser, "Automatic detection, segmentation and assessment of snoring from ambient acoustic data," *Physiol Meas,* vol. 27, no. 10, pp. 1047-1056, 2006.

[6] A. Karunajeewa, U. Abeyratne and C. Hukins, "Silence-breathing snore classification from snore related sounds," *Physiol Meas,* vol. 29, no. 2, pp. 227-243, 2008.

[7] R. Jane, J. Fiz and J. Sola-Soler, "Automatic Snoring Signal Analysis in Sleep Studies," in *IEEE EMBS*, Cancun, Mexico, 2003.

[8] A. K. Ng, T. S. K. E and P. Kathiravelu, "Snore Signal Enhancement and Activity Detection," *EEE Transactions on Biomedical Engineering,* vol. 55, no. 10, pp. 2332-42, 2008.

[9] H. Kreivi, T. Salmi, P. Maasilta and A. Bachour, "Screening of snoring with an MP3 recorder," *Sleep Breath,* vol. 17, no. 1, pp. 77-84, 2013.

[10] GsMarena, "Vodafone V860 Specs," 2011. [Online]. Available: http://www.gsmarena.com/vodafone_v860_smart_ii-4869.php. [Accessed October 2013].

[11] NLL, "Google Play Store," 2013. [Online]. Available: https://play.google.com/store/apps/details?id=com.nll.asr&hl=en. [Accessed 1 October 2013].

[12] A. Zaffaroni and H. Conor, "SleepMinder: An Innovative Contact-Free Device for the," Minnesota, 2009..

[13] T. Giannakopoulos, "A method for silence removal and segmentation of speech signals, implemented in Matlab," Report presented in Mathworks File Exchange Website, 2010.

[14] Q. Gu, Z. Li and J. Han, "Generalized Fisher Score for Feature Selection," Champaign, 2012.