An Inhaler Tracking System Based on Acoustic Analysis: Hardware and Software

Wenyang Xie, *Student Member,* Patrick Gaydecki, *Senior Member,* Ann-Louise Caress

Abstract—In treating asthma and chronic obstructive pulmonary disorder (COPD), acquisition of authentic and effective feedback from patients on regimen adherence is difficult. Face-to-face and oral reporting methods do not satisfy current intelligent medication best practices. This paper presents a system to track and analyse daily inhaler usage. A portable electronic device that attaches to the inhaler uses an accelerometer and capacitive sensors to detect users' motion and an embedded digital microphone to capture sounds while the inhaler is in use. In terms of analysis, sound features are extracted, and breath phases are identified by employing a hidden Markov model (HMM) with a Gaussian mixture model (GMM). A feature template is also constructed and used to search for and identify 'canister pressed' events. The system provides objective feedback, quantifying asthma and COPD patients' adherence to medication regimens. Although interest in asthma adherence to medication regimens is growing, there is still a relative paucity of research and, indeed, compliance devices in this area; the tracking system can help doctors better understand the patient's condition and choose an appropriated treatment plan. At the same time, patients can also improve their self-management by system feedback.

Index Terms—Acoustic monitoring, breath phase identification, inhaler techniques, HMM-GMM, RF, SVM

I. INTRODUCTION

THE number of patients with asthma and chronic obstructive pulmonary disease (COPD) is gradually increasing. A recent projection predicts COPD will be the third leading cause of death in much of the world by 2020 [1]. Typical treatment for asthma and COPD is inhaled medications, but this requires good co-ordination of exhalation and inhalation, especially when using the metered dose inhaler (MDI) technique [2]. Appropriate use of the MDI technique can deliver the correct dose of medicine into patients' lungs; however, one study showed that only 14% of patients with asthma used their MDI correctly, while accurate use for COPD patients was 15% [3]. To achieve therapeutic benefits, it is critical that patients use inhaler techniques correctly and that doctors and health providers understand patients' conditions and the necessary treatment. Currently, the ability to assess how regularly and effectively patients use their inhaler on a daily basis is limited. Such assessment of patients' inhaler technique is usually done face to face by a health provider, which is, in practice, highly inefficient. Some inhalers are designed to provide time of use as well as dose count, but few can provide any indication of whether doses were inhaled correctly or at all. Few devices allow accurate and objective assessment of how patients use the full range of inhalers on a daily basis during routine care [3-5]. This low rate of accurate use and the difficulty of confirming adherence implies that patients are often not compliant with health providers' or doctors' recommendations. Incorrect use of inhaler techniques is a significant factor in prescribed treatment [6]. Moreover, poor inhaler use also unnecessarily increases medical budgets and pressure on healthcare organizations, while patients suffer from inefficacy of treatment and other potential problems. Currently, dose counting is used to track patients' inhaler use in clinical trials [7], but only a limited number of inhalers have in-built dose counters (usually they are expensive; these are not inhalers that would routinely be prescribed in practice). Furthermore, tracking daily inhaler use is still difficult due to the lack of a reliable method [4]. Thus, there is a considerable space for improvement in assessing whether patients are taking the right number of doses. It is clearly important, given the rate of nonadherence in asthma and COPD patients and the known adverse, potentially life-threatening, consequences of lack of adherence.

An algorithm to identify breath sounds is also significant to the implementation of adherence analysis. One study used power spectral density (PSD) to discriminate between inhalation and exhalation, with an accuracy of 92.8% in a relatively silent environment [8]. Discrimination of inhalation and exhalation is based on a single segment of breath sound. However, a serial raw-audio signal not only contains multiple segments of interest but also contains unexpected information, such as voice, music, and noise. It is thus necessary to demarcate breath sounds within a raw audio signal. One study detected inhalation sounds based on Mel-frequency cepstral coefficients (MFCC), with an accuracy of 89% [9]. Ruinskiy and Lavner constructed an efficient algorithm to label breath sounds in audio signals, with a correct identification rate of 98% [10]. Two studies used similar methods, template matching, to detect breath sounds. The template was constructed by using singular value vectors, which used singular-value decomposition (SVD) to degrade the matrix of MFCC, to represent the essential characteristics, and to increase efficiency

[9],[10]. Various methods are used for classification in analysis of acoustic sounds, such as artificial neural networks (ANN), k-nearest neighbors (KNN), the Gaussian mixture model (GMM), the hidden Markov model (HMM) and Time delay neural network (TDNN) [11][12]. For example, GMM was employed to classify acoustic signal of moving targets, its classification accuracy was up to 92.15% in mono input [13]. In practice, the duration of inhalation and exhalation is unpredictable. The variable length of each segment cannot be represented by fixed-size matrices. Therefore, classification methods are required to process the variably sized sequence.

This paper presents a system including hardware and software that tracks inhaler usage based on acoustic monitoring. The electronic device has low-power consumption, is miniaturised, flexible and can be retrospectively fitted to standard inhalers. The software uses efficient, embedded algorithms to detect and recognize inhaler use. The original sounds are truncated based on sound activity detection, with the aim of removing unnecessary silent parts. HMM-GMM, which is one of the most popular solutions in sequential prediction [14], is used to create an acoustic model and identify breath phases. The Gaussian mixture model (GMM) is used to model the feature distribution and can be thought of as a single state HMM. Usually, the GMM contains several mixtures to fit the input signal and then its mixture distribution is used as the emission probability of an HMM. The GMM is particularly appropriate in this context due to the variable nature of the signal being classified. Dose counting is also based on audio analysis, with basic audio features classified by a random forest and support vector machine (SVM) approaches.

II. System Design

To measure the rates of accurate inhaler use, this system is designed to monitor daily inhaler use and analyse the collected data. It provides assessments that show whether patients are in compliance with the recommended usage steps. The system has two main parts. The hardware implements automatic capturing and recording of breath sounds, while the software identifies breath phase and analyses adherence on a PC. The recommended steps of inhaler use (see Table I) were developed according to Asthma UK [15].

Three of the steps in the recommended instructions for MDI use are most significant and require good co-ordination of

TABLE I
RECOMMENDED INHALER USE STEPS

MDI Step by step	Phases
1. Shake the inhaler well before use	-
2. Remove the cap	-
3. Empty lungs of air before inhalation	Exhalation
4. Place inhaler in mouth between teeth and close mouth around it.	-
5. Inhale slowly, at the same time pressing the top of the inhaler once while continuing to breathe in slowly until a full breath is taken.	Inhalation & Spray
6. Remove inhaler from mouth, hold breath around 10 seconds and breathe out	Exhalation

breath while pressing the canister. Therefore, the designed system focuses on tracking and timing of the breath phase. A system block diagram is shown in Fig. 1.

The *exhale slowly phase* has the lowest accuracy rate for all recommended steps. If this stage of the procedure is disregarded, accurate identification rises to 96.9% [16], which means that the exhalation step accounts for a large proportion of the failure to adhere to the proper technique. Therefore, the monitoring device must record the entire process of inhaler use. Tracking should be triggered, and the record-function should be run before the first exhalation. This triggering mechanism is implemented by a combination of detection of motion and touch. Furthermore, it also activates the system from standby status, which is a mode used to extend battery duration, reduce frequency of recharging, and minimise inconvenience of daily use. The triggered device then records the whole breathing progress and limits the frequency spectrum within a specific range via digital filters. Finally, raw audio data is labelled with a time stamp and saved to storage memory. Adherence is analysed on a PC, as shown in the right part of Fig. 1. It contains four parts, including data pre-processing (such as normalisation, hamming framing, and feature extraction), model training, construction of the 'pressing event' template, and fusion of the results. All sound data are split into segments. Each truncated part is then converted to a feature matrix of Mel-frequency cepstral coefficients (MFCC) and normalised onto a common metric for further training. As for the canister pressed event detection, a fixed sliding window scans the whole
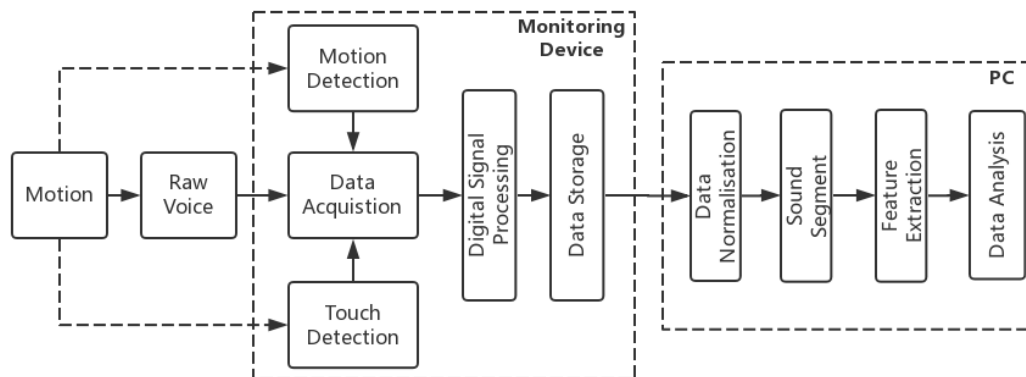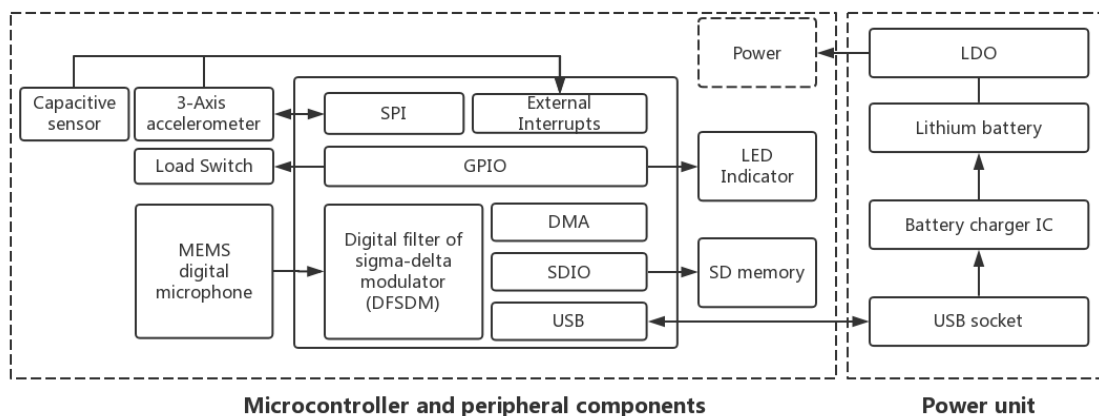

Fig. 1. System block diagram.

Fig. 2. Hardware block diagram.

recording and extracts audio features, such as MFCC, zero-crossing rate, and spectral centroid. In the final stage, the trained classifier outputs the predicted result.

A. Hardware Architecture

A diagram of the hardware is shown in Fig. 2. Its purpose is to capture the user's motion, record breath sounds and store data. The electronics feature low power consumption since the microcontroller is sleeping in standby mode most of the time. For reduction of static power consumption, this mode also turns off most of the peripheral components via a load switch chip. The cooperative triggering of the capacitive sensor and 3-axis accelerometer activates the microcontroller when the device is picked up and prepared for use. The activated microcontroller then switches on all peripheral loads, such as LED indicator, SD card, and MEMS microphone. According to empirical research on usage duration, the average recording is 25 seconds and the external memory stores this time-stamped audio. Finally, the microcontroller turns off the unnecessary peripherals and returns to standby mode until the next activation (program structure is shown in Fig. 3). The size comparison given in Fig. 4 shows that the electronics are very small and can attach to various existing inhalers with specific enclosures.
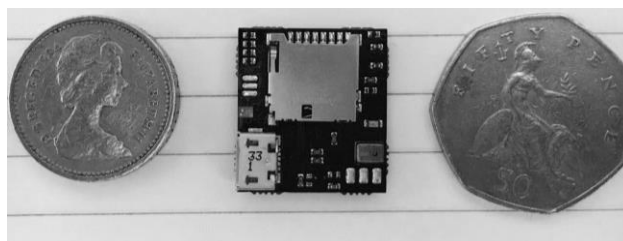

Fig. 3. Microcontroller program structure.

Fig. 4. Monitoring electronics.

B. Algorithm Development

The inhalation and exhalation phases are characterized by significantly different spectral features, as shown by the example in Fig. 5. The spectrum of the inhalation phase is typically broad band in nature, extending from 20 to 8000 Hz. In contrast, the energy of the exhalation signal is maximal around 500 Hz and gradually falls away from this point.
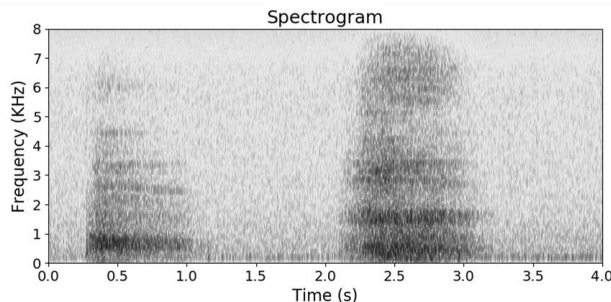

Fig. 5. Breath sounds spectrogram (left: exhalation; right: inhalation).

Therefore, a method was developed for the automatic detection of breath sound in speech and identification of the breath phases. The algorithm consists of two main parts. The first is detection of the breath phases, which is used to train an acoustic model via supervised learning. First, during pre-processing, the raw signal is limited to filter out unexpected noise within the range 20 Hz to 8 kHz. Sound activity detection (SAD) is employed for slicing the temporal sequence while retaining the appropriate segments for the subsequent computation of the MFCC. In the next step, the labelled training samples are split into overlapping frames. A single frame is computed as an MFCC feature vector, normalised and the processed matrix trained via the GMM-HMM method; this yields an acoustic model. The test samples are used for
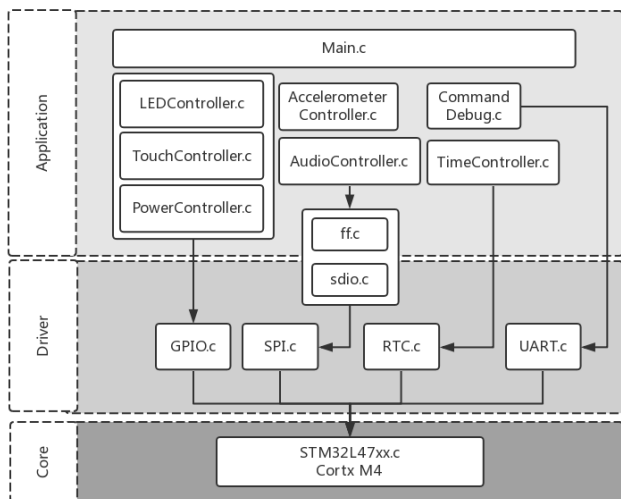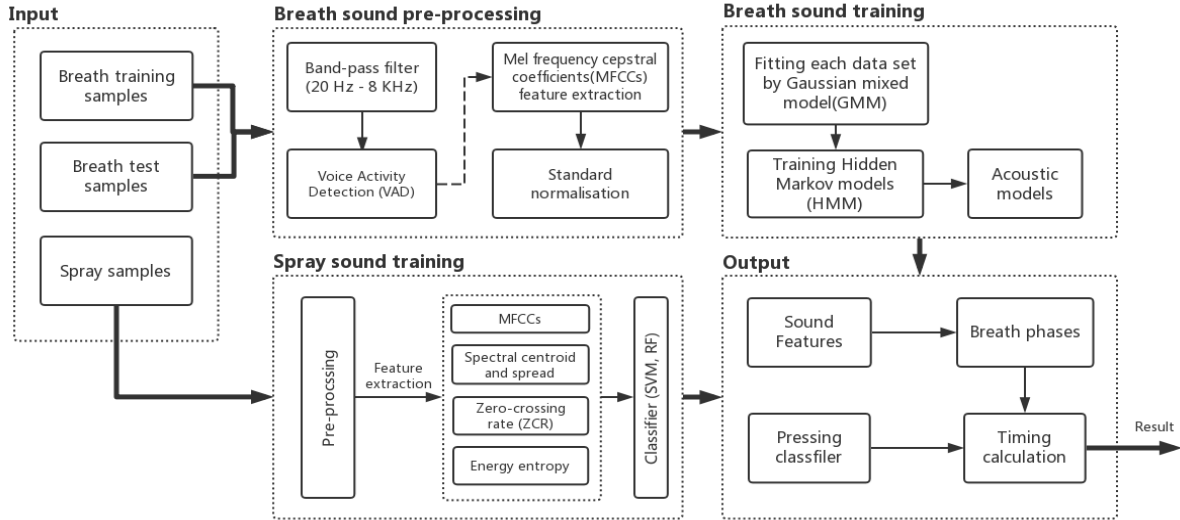
Fig. 6. Flow diagram of the algorithm.

performance validation. Additionally, to quickly detect when the canister is pressed, a template is built. The final output is the merged results of the two parts, calculating the time intervals between identified locations. The whole process is shown in Fig. 6.

1) *Breath sound pre-processing*
- *Band-pass filter*
  The raw signal passes through a Butterworth band-pass filter (BPF), which has a pass-band between 20 Hz and 8 kHz, to remove noise and enhance the accuracy. Outside of this band, analysis revealed that there is almost no contribution from the breath sounds. Attenuation of the out-of-band signal, which derived from a number of ambient sources (often percussive in nature) was therefore essential in maximising the signal-to-noise ratio.
- *Sound activity detection*
  The implementation of SAD reduces the computation load and increases the efficiency of detection. It is based on the crest factor, expressed in decibels as

$$Crest\_factor_{dB} = 20*\log_{10}\frac{\left|f_{peak}(x)\right|}{f_{RMS}(x)} \quad (1)$$

$$f_{RMS}(x) = \sqrt{\frac{\sum_{i=0}^{N-1} x_i^2}{N}}$$

The crest factor reflects the ratio of the peak and mean levels and indicates the activity of sound when that ratio exceeds a given threshold. Fig. 7 illustrates the relationship between the active sounds and the crest factor. Each margin of detection window is 10% of the raw sound, to maximise the retention of useful information.
- *MFCC extraction*
  The feature extraction block separately divides the effective segments into overlapping frames; the

window length and overlap between windows are 25 ms and 10 ms, respectively. Each of the MFCC features is computed and generates a vector. This process is shown in Fig. 8.
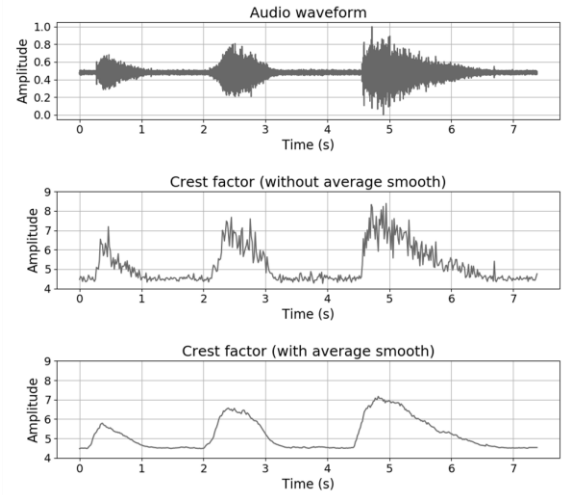

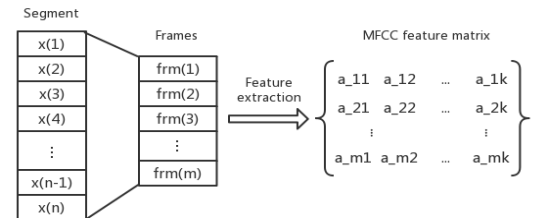Fig. 7. Sound activity detection by short-term crest factor.

Fig. 8. MFCC feature matrix.

- *Standard normalization*
  To avoid mismatched metrics in the training phase, it is necessary to map all data on the same scale of statistics. The function is given by

$$f_{normal}(x) = \frac{x - \mu}{\sigma} \quad (2)$$

where μ is the mean of the signal and σ is the standard deviation.

2) Breath sound training

A multivariate Gaussian model was employed to express the probability distribution of the feature matrix. This is given by [17]

$$p(x) = \sum_{k=1}^{K} \pi_k N\left(x \mid \mu_k, \Sigma_k\right)$$
$$= \sum_{k=1}^{K} \frac{\pi_k}{(2\pi)^{\frac{D}{2}} |\Sigma_k|^{\frac{1}{2}}} e^{\left[-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1}(x-\mu_k)\right]}, (\pi_k > 0) \quad (3)$$

Where mixture weights π, co-variance matrix Σ, and mean μ are estimated via the expectation-maximization (EM) algorithm, which aims to find the most appropriate parameters for a given fit.

The computed GMM is considered as a single and special state of HMM [18]. The emission probability of the latent state, in fact, is calculated by the previous step as

$$b_j(x) = \sum_{k=1}^{K} \pi_k N\left(x \mid \mu_{jk}, \Sigma_{jk}\right), (\pi_k > 0) \quad (4)$$

These observations *X* based on the output of states (such as inhalation, exhalation, and non-breath) correspond to the computed GMM distributions in the previous step and are shown in Fig. 9.

A_{ij} is transition probability. According to the Bayesian theorem, the posterior probability is expressed as follows:

$$p(S_j \mid X, \theta) = \frac{p(X, S_j \mid \theta) p(S_j \mid \theta)}{p(X \mid \theta)}, \quad (5)$$

$$\theta = \{A, \mu_j, \Sigma_j\}$$

Using the conditional independence property [18], the expression takes the form

$$p(S_j \mid X, \theta) = \frac{a_j(n)\beta_j(n)}{p(X \mid \theta)},$$

$$a_j(n) \equiv p(x_1, ..., x_n, S_j \mid \theta), \quad (6)$$

$$\beta_j(n) \equiv p(x_{n+1}, ..., x_N \mid S_j, \theta)$$

Assuming the observed sequence *X* is known, the Baum-Welch algorithm can be employed to evaluate (6) to obtain HMM parameters.
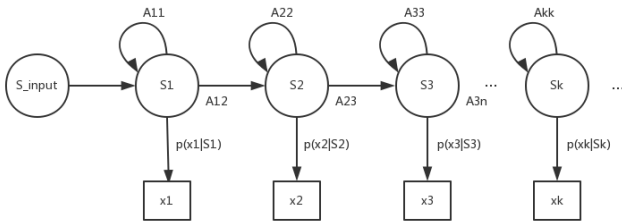
Fig. 9. Latent variables with observed variables.

3) Detection of spraying

Detection of effective usage (dose counting) is indispensable to recording the frequency and duration of medical usage of MDIs. It also reflects the patients' dependency on the prescription, i.e. whether patients are decreasing or increasing their use. Therefore, detection must be effective and efficient.

In terms of MDI, the acoustic signal arises from the ejection of an aerosol cloud, by a propellant, from the metering valve [19], which then expands in an expansion chamber and is sprayed out by an actuator nozzle. Analysis of sample recordings shows a typical pressing feature indicated by a peak in the frequency domain within the range of 1600 Hz to 1800 Hz. Fig. 10 shows that the majority of the energy is aggregated on this sub-band and the rest is widely spread on the whole band. Furthermore, the graph of autocorrelation, shown in Fig. 10, displays periodic characteristics. Furthermore, it is hard to extract a fixed central frequency to identify the sound, because of changes in the container pressure and pressing jitter. Fig. 11 shows the major frequency statistics of 46 samples; the average frequency is 1658 Hz, and the standard deviation is 37.7. Therefore, to avoid unnecessary computation, a band gate is set within 1658 Hz ±10% to block any unsatisfactory signal, before the feature extraction process. Capturing the major frequency of the sound allows partial identification of the unknown signal, but only considering this will result in low specifity. Addtional features are considered to improve performance and accuracy of detection.
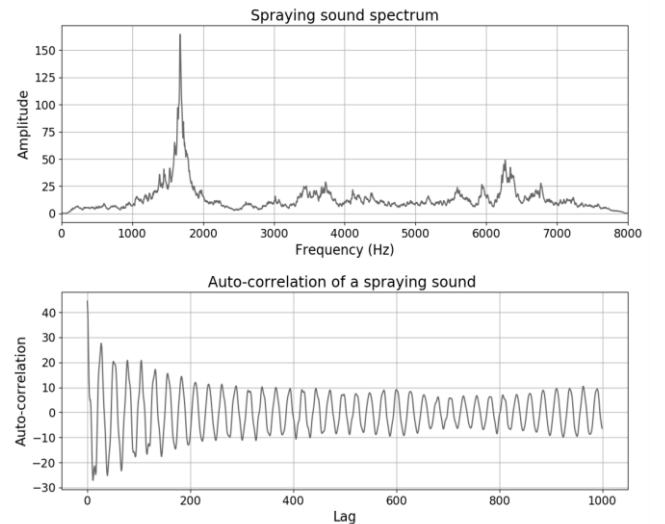
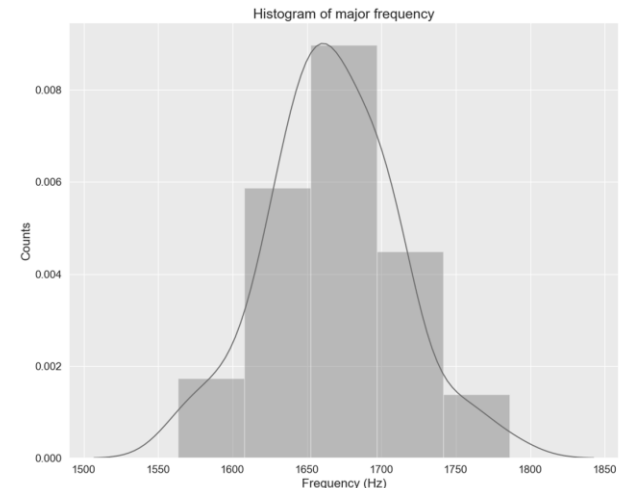Fig. 10. Spectrum (up) and enhanced autocorrelation (down) of spraying sound.

Fig. 11. Statistics of major frequency.

- *Extraction of acoustic features*
  a) MFCC is an appropriate feature to reflect hearing perception. Equation (7) normalises each coefficient and averages them:

$$M_{norm} = \frac{1}{N}\sum_{j=1}^{N}\frac{M_j^i - mean(M_j^i)}{var(M_j^i)} \qquad (7)$$

  where M^i is the ith coefficient in the MFCC matrix and M_j is the jth sub-frame.

  b) The zero-crossing rate is an important feature in this detection. It is expressed by

$$f_{zcr} = \frac{1}{N}\sum_{n=0}^{N-1}\left(\text{sgn}(x[n]) - \text{sgn}(x[n-1])\right) \qquad (8)$$

  c) The spectral centroid indicates the locations where most of the energy is concentrated. The brightness (in a timbral sense) of a sound depends on this. It is expressed as the spectral-weighted sum divided by its unweighted sum [20]:

$$f_{ct}(n) = (\sum_{k=0}^{\frac{N}{2}-1}k\left|F(k,n)\right|)/(\sum_{k=0}^{\frac{N}{2}-1}\left|F(k,n)\right|) \qquad (9)$$

  where N is the fast Fourier length of each temporal block.

  d) The spectral spread describes the flatness of the spectrum; it is an indicator of the noisiness [20]. This feature is given by

$$f_{ss}(n) = \left[\frac{(\sum_{k=0}^{\frac{N}{2}-1}(k - f_{ct}(n))\cdot k\left|F(k,n)\right|)}{(\sum_{k=0}^{\frac{N}{2}-1}\left|F(k,n)\right|)}\right]^{\frac{1}{2}} \qquad (10)$$

  e) Energy entropy is the measurement of the distribution of energy. It is based on Shannon information entropy. Increment of energy entropy means the sound contains more uncertainty of

energy and vice versa. It is given by [21]

$$H(x) = E\{I(x_i)\} = \sum_{i=1}^{N} p_i \log_2(p_i) \qquad (11)$$

where p is the probability of obtaining useful information. According to this concept, the information probability can be substituted by the ratios of sub-blocks energy to the total energy. Therefore, energy entropy is defined as

$$H(n) = \sum_{i=0}^{M}\frac{\left|f_i(n)\right|^2}{\left|f(n)\right|^2}\log_2(\frac{\left|f_i(n)\right|^2}{\left|f(n)\right|^2}),(M \le N) \qquad (12)$$

- *Classification method*

The classification step uses two popular classifiers on a small dataset. Random forest (RF) is a collection of 'weak' classifiers (decision tree) which are combined and contribute to a strong classifier [22]. It stochastically selects subset of features and generates serval mutually independent trees. The output prediction is voted by the 'weak' classifiers [23]. The second classifier is SVM, which is a discriminative method that is implemented in regression and classification. Its classification strategy is based upon choosing an appropriate hyperplane that contains the maximum decision boundaries [18], [24], [25]. Recently, SVM has gained popularity as a classification method in areas of medical diagnosis and detection [27-30].The hypothesis parameters of classifiers are tuned via a grid search function [31]. Fig. 12 shows the visual result for the continuous pressing test. The first row of the figure shows the SVM prediction and the bottom row shows the results based on RF classification. In Fig. 12, SVM and RF both require approximately 0.012 seconds for prediction of a segment fifteen seconds in length, thus showing similar performance with this small number of features.
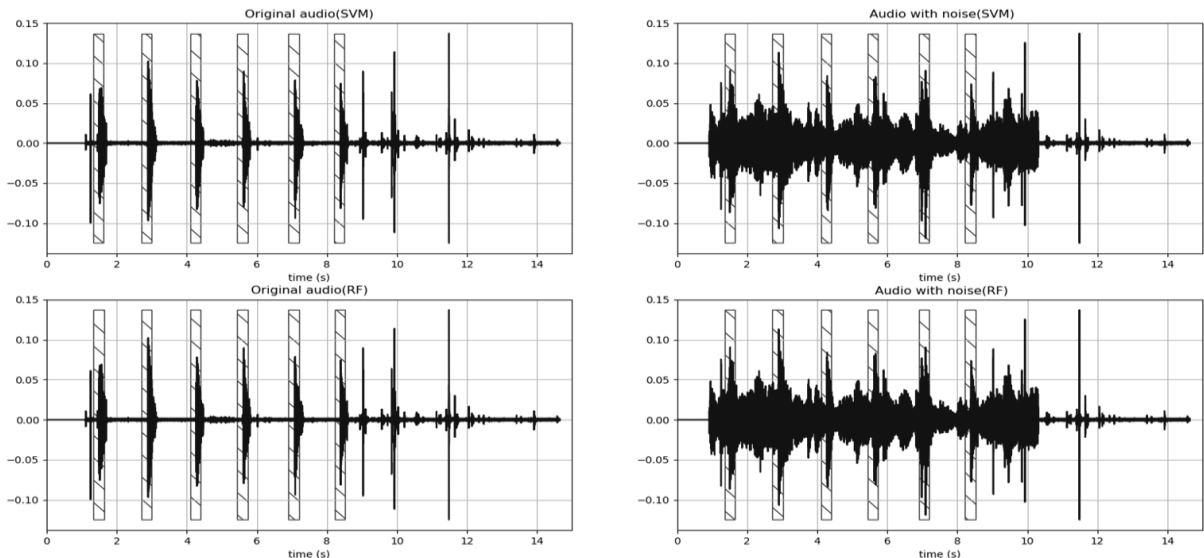

Fig. 12. Left: detections of pure audio; right: detections of audio in a noisy environment.
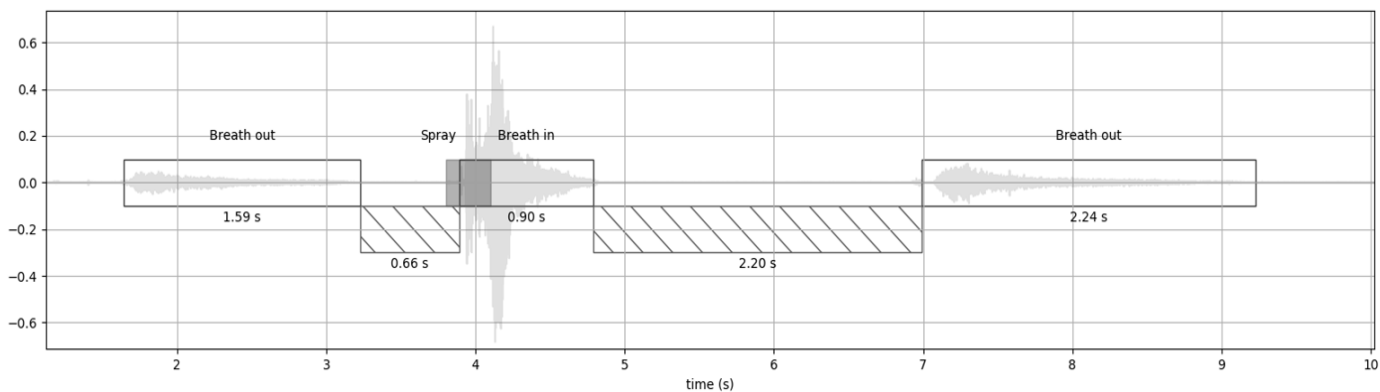
Fig. 13. Analysis of adherence to proper usage.

4) Detection output

The final output combines all the detected events. Fig. 13 displays a waveform of actual inhaler use, in which the different steps are automatically labelled by the detection algorithms described in this paper. The transparent rectangles mark the locations of the breath timeline, while the dark square is the detected spray event and the time intervals within the steps are shown by hatched rectangles. This format provides useful information, such as the sequence of use steps, the duration of each step, and the number of doses. For example, Fig. 13 shows that the user held their breath for only a short time of a few seconds, which may cause the medicine to be immediately exhaled [32].

III. PERFORMANCE EVALUATION

A. Device-Triggered Evaluation

A study was designed to evaluate the reliability of the electronics and the collection of information, corresponding to the ease of use of the smart inhaler. The study took place over two days and included four volunteers. Before the study, the volunteers were given minimal training on the correct use of the device, and then they were asked to count the number of times they used the monitoring device and record it as a written record. Table II shows that a total of 62 uses were reported by volunteers, but due to false triggering, the devices recorded 65 audio traces containing 3 non-related records. After filtering out the incorrectly triggered recordings, the accuracy increased to 93.5%.

The electronics developed and used to acquire volunteers' data is shown fitted to an inhaler in Fig. 14. At the time of writing, this inhaler is the most widely used design. As our system evolves and becomes more reliable, it is planned to trial

it on a number of different inhaler models. Since the accuracy and reliability of the system is based upon an automated algorithm that identifies the various phases of the signal, we do not anticipate that the change of inhaler will affects the system's overall performance.

Fig. 14. Experimental device.

B. Breath Phase Detection

In this test, healthy volunteers were divided by gender into two groups; in each group, one volunteer was trained in or had previous experience using the inhaler. Group 1 recorded a total of 60 breath sounds each, and group 2 recorded 20 uses for each volunteer. The test was carried out in an indoor environment. Table III shows that inhaler use of volunteers A, C, D and F, who had related experience, was correctly recognised at a rate of over 90%. However, for volunteers B and E, who were both inexperienced participants, the average accuracy fell to 88%. As a result of unskilled operation, the breathing sounds are quieter and inhalation is slow. This results in lower accuracy of detection for volunteers B and D. Although more robust noise rejection algorithms are still being developed, preliminary tests to establish performance in the presence of noise has been assessed, as evidenced in Table III. The original high quality signals were degraded by adding randomly selected noise data, which included traffic and urban street noise. As shown in the final column, with an SNR of 25 dB, the accuracy of the system drops to a mean of 79.77%, in comparison to a mean of 92.83% under ideal conditions. The reduction of performance in the presence of noise is inevitable, but will be ameliorated by compensation algorithms currently under development.

TABLE II
SUMMARY OF VOLUNTEER RECORDINGS

Volunteers	Use times	Recorded samples	Correct recordings	Audio duration
A	15	16	14	25s
B	7	7	6	25s
C	25	25	25	25s
D	15	17	13	25s

TABLE III
SUMMARY OF RECOGNITION PERFORMANCE

Volunteers	Gender	Recorded	Trained or used[1]	Accuracy % (Without noise)	SNR 25 (dB)
A	male	30	Y	96.5%	83.5%
B	male	30	N	87.9%	76.1%
C	male	45	Y	95.8%	82.8%
D	female	20	Y	94.3%	80.2%
E	female	20	N	88.1%	74.8%
F	female	45	Y	94.4%	81.2%

1. Y - yes; N - no

TABLE IV
PERFORMANCE OF CLASSIFIERS – SPRAY IDENTIFICATION

Classifiers	Precision	Recall	F1	Accuracy
SVM	1.000	0.978	0.989	0.985
RF	0.967	0.989	0.978	0.969

C. Spray Identification

A total 130 samples were collected in an indoor environment, among which, 90 contained spray sounds and while the other 40 were irrelevant noises. Table IV shows a comparison of the performance of the two classifiers. After optimising the hyper-parameters for the two classifiers, RF results in higher recall than SVM, indicating that RF has strong discriminative ability for the relevant sample. However, the precision of RF is slightly lower than that of SVM. In this case, it is difficult to judge the performance of the two classifiers in this dataset. Therefore, a combinative average of the precision and recall was used as an indicator - F1 score. According to the F1 score, SVM shows better performance than RF, and its accuracy reaches 96.1%.

IV.   CONCLUSION

Tracking and quantifying the adherence of patients to prescribed medication regimes is an important factor in increasing the efficacy of treatment, especially with the rapid development of medicine technology. This paper presents a monitoring system that includes both hardware and software that provides a practical method for estimating the adherence of pMDI users to medication regimens. Although other methods are possible, based on artificial intelligence or machine learning, the algorithm presented here was efficient and robust, even in the presence of some noise. However, the system also needs improvement in future work. Detection is compromised in very noisy surroundings, such as on a busy street or in shopping malls; its accuracy degrades in such environments. Therefore, a noise estimation model for such common situations and breath sound enhancements are the next focus of research. Furthermore, the entire training dataset was collected in a quiet environment. In terms of signal analysis, a pure signal can reflect the characteristics of the target. However, this may cause the classifier to become incapable of generalizing, which can be considered as an example of an overfitting problem. Therefore, the collection of positive samples with different

backgrounds is also a focus of future research. Initial tests suggest that the system has significant potential for the subjective monitoring of patients, aid clinicians in the analysis of compliance, and will ultimately improve medical self-management for users of inhalers.

REFERENCES

[1]   P. Guide and T. O. Copd, "Global Initiative for Chronic Obstructive Lung A Guide for Health Care Professionals Global Initiative for Chronic Obstructive Disease," *Glob. Initiat. chronic Obstr. lung Dis.*, vol. 22, no. 4, pp. 1–30, 2017.

[2]   R. D. Restrepo *et al.*, "Medication adherence issues in patients treated for COPD," *Int. J. COPD*, vol. 3, no. 3, pp. 371–384, 2008.

[3]   V. Press *et al.*, "Misuse of Respiratory Inhalers in Hospitalized Patients with Asthma or COPD," *J. Gen. Intern. Med.*, vol. 26, no. 6, pp. 635–642, 2011.

[4]   S. D'Arcy *et al.*, "A method to assess adherence in inhaler use through analysis of acoustic recordings of inhaler events," *PLoS One*, vol. 9, no. 6, pp. 1–11, 2014.

[5]   M. S. Holmes, S. D'Arcy, R. W. Costello, and R. B. Reilly, "Acoustic Analysis of Inhaler Sounds From Community-Dwelling Asthmatic Patients for Automatic Assessment of Adherence," *Transl. Eng. Heal. Med. IEEE J.*, vol. 2, no. 1, pp. 1–10, 2014.

[6]   J. T. Lindsay and L. G. Heaney, "Nonadherence in difficult asthma - facts, myths, and a time to act.," *Patient Prefer. Adherence*, vol. 7, pp. 329–36, 2013.

[7]   S. P. Peters *et al.*, "Tiotropium Bromide Step-Up Therapy for Adults with Uncontrolled Asthma," *N. Engl. J. Med.*, vol. 363, no. 18, pp. 1715–1726, 2010.

[8]   M. S. Holmes, S. D'Arcy, R. W. Costello, and R. B. Reilly, "An acoustic method of automatically evaluating patient inhaler technique," *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2013, pp. 1322–1325, 2013.

[9]   M. S. Holmes *et al.*, "Automatic identification and accurate temporal detection of inhalations in asthma inhaler recordings," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 2595–2598, 2012.

[10]   D. Ruinskiy and Y. Lavner, "An effective algorithm for automatic detection and exact demarcation of breath sounds in speech and song signals," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 15, no. 3, pp. 838–850, 2007.

[11]   R. Palaniappan, K. Sundaraj, and N. U. Ahamed, "Machine learning in lung sound analysis: A systematic review," *Biocybern. Biomed. Eng.*, vol. 33, no. 3, pp. 129–135, 2013.

[12]   A. R. Abu-el-quran, R. A. Goubran, and A. D. C. Chan, "Security Monitoring Using Microphone Arrays and Audio Classification," *Ieee Trans. Instrum. Meas.*, vol. 55, no. 4, pp. 1025–1032, 2006.

[13]   J. Huang, X. Zhang, F. Guo, Q. Zhou, H. Liu, and B. Li, "Design of an Acoustic Target Classification

System Based on Small-Aperture Microphone Array," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 7, pp. 2035–2043, 2015.

[14] M. J.~Beal, Z. Ghahramani, and C. E. Rasmussen, "The Infinite Hidden Markov Model," *Adv. Neural Inf. Process. Syst. 14*, pp. 577–584, 2002.

[15] Asthma UK, "Using your inhalers," *www.asthma.org.uk*, 2017. [Online]. Available: https://www.asthma.org.uk/advice/inhalers-medicines-treatments/using-inhalers/#Howdoesgoodinhalertechnique... [Accessed: 16-Nov-2017].

[16] H. Lee, S. Boo, Y. Lim, S. Kim, and I.-A. Kim, "Accuracy of inhaler use in patients with chronic obstructive pulmonary disease," *Clin. Nurs. Res.*, vol. 23, no. 5, p. 560, 2014.

[17] P. R. Prem and A. R. Murthy, "Acoustic emission monitoring of reinforced concrete beams subjected to four-point-bending," *Appl. Acoust.*, vol. 117, pp. 28–38, 2017.

[18] C. M. C. C. M. Bishop, "Pattern Recognition and Machine Learning," *J. Electron. Imaging*, vol. 16, no. 4, p. 049901, 2007.

[19] T. Noakes, "Medical aerosol propellants," *J. Fluor. Chem.*, vol. 118, no. 1–2, pp. 35–45, 2002.

[20] B. Sturm, *Book review: An Introduction to Audio Content Analysis: Applications in Signal Processing and Music Informatics by Alexander Lerch*. Hoboken, N.J.: Hoboken, New Jersey : Wiley,; [Piscataqay, New Jersey] : IEEE Xplore, 2013.

[21] M. Borda, *Fundamentals in Information Theory and Coding*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2011.

[22] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[23] P. Jiang, H. Wu, W. Wang, W. Ma, X. Sun, and Z. Lu, "MiPred: Classification of real and pseudo microRNA precursors using random forest prediction model with combined features," *Nucleic Acids Res.*, vol. 35, no. SUPPL.2, pp. 339–344, 2007.

[24] D. Wilk-Kolodziejczyk, K. Regulski, and G. Gumienny, "Comparative analysis of the properties of the nodular cast iron with carbides and the austempered ductile iron with use of the machine learning and the support vector machine," *Int. J. Adv. Manuf. Technol.*, vol. 87, no. 1–4, pp. 1077–1093, 2016.

[25] Y. Huang, H. Yang, X. Qi, R. Malekian, O. Pfeiffer, and Z. Li, "A novel selection method of seismic attributes based on gray relational degree and support vector machine," pp. 1–16, 2018.

[26] Y. Zhang, P. Wang, T. Ni, P. Cheng, and S. Lei, "Wind power prediction based on LS-SVM model with error correction," *Adv. Electr. Comput. Eng.*, vol. 17, no. 1, pp. 3–8, 2017.

[27] K. Kancherla and S. Mukkamala, "Feature selection for lung cancer detection using SVM based recursive feature elimination method," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7246. pp. 168–176, 2012.

[28] C. Désir, C. Petitjean, L. Heutte, L. Thiberville, and M. Salaün, "An SVM-based distal lung image classification using texture descriptors," *Comput. Med. Imaging Graph.*, vol. 36, no. 4, pp. 264–270, 2012.

[29] W.-F. Wang, C.-Y. Yang, and Y.-F. Wu, "SVM-based classification method to identify alcohol consumption using ECG and PPG monitoring," *Pers. Ubiquitous Comput.*, vol. 22, no. 2, pp. 275–287, 2018.

[30] Y. Zhang, X. Ji, and S. Zhang, "An approach to EEG-based emotion recognition using combined feature extraction method," *Neurosci. Lett.*, vol. 633, pp. 152–157, 2016.

[31] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2012.

[32] Asthma UK, "Using your inhalers." [Online]. Available: https://www.asthma.org.uk/advice/inhalers-medicines-treatments/using-inhalers/. [Accessed: 09-May-2018].

Wenyang Xie was born in Chongqing, China, in 1991. He received the B.S. degree in electronic information engineering from Chongqing University of Posts and Telecommunications, Chongqing, China, in 2014. He is currently pursuing the Ph.D. degree in electrical and electronic engineering with the University of Manchester, Manchester, UK.

His research interests include signal processing, machine learning, data analysis, Internet of things (IoT), and embedded system.

Patrick Gaydecki, PhD, FIET, SMIEEE, FInstP, MInstNDT, CPhys, CEng is Professor of Digital Signal Processing (DSP) in the School of Electrical and Electronic Engineering at the University of Manchester. With his team, he develops systems and software in the fields of audio digital signal processing, wearable devices, IoT and mobile systems, biomedical signal processing and instrumentation. He has authored a book on DSP and has published over 200 journal and conference papers. His research has been funded by EPSRC, Innovate UK, the DTI, the BHA, ESRC, MIMIT, TWI, The HA, and many industrial sponsors. He is the inventor of a range of commercial real-time DSP products for audio signal enhancement, including Signal Wizard 2.5, Signal Wizard 3, vsound, dCello and Soundtrack. He is also the co-founder of Signal Wizard Systems Ltd, whose customers include private individuals and organizations such as Disney, London Central Communications Ltd, The US Woods Hole Oceanographic Institution, Pure Technologies Canada, TeliaSonera, Canadian Defence Research and Development and US Government agencies.

Ann Caress graduated from the University of Manchester in 1986. Her research career began in 1988, as a research associate on a project developing computer assisted learning for people with kidney failure. Ann undertook her doctoral studies at the University of Liverpool (PhD awarded 1996).

After receiving her Ph.D. degree, Ann became a lecturer at the University of Liverpool. She returned to the University of Manchester as a lecturer, becoming a Professor in 2009. In 2018, she joined the University of Huddersfield, taking on the role of Professor of Health Services Research. She has held two NHS joint appointments, at the Christie Hospital NHS Foundation (2001-2003) and Manchester University NHSFT (formerly UHSM) from 2010-2018. Ann's primary research interests relate to living with and management of long-term health problems, primarily in adults and mainly in the renal and respiratory fields. She has particular interests in self-management, patient and carer experience and novel approaches to service delivery. She is a strong advocate of patient and public involvement in service delivery and research and held executive roles in INVOLVE from 1997-2013. Ann is extremely active in promoting research nationally and internationally, with research collaborations in such countries as Switzerland, Hong Kong and Canada.