SCIENCE CHINA
Information Sciences

• MOOP •

March 2021, Vol. 64 134101:1–134101:3
https://doi.org/10.1007/s11432-019-1510-6

Using breath sound data to detect intraoperative respiratory depression in non-intubated anesthesia

Sikai WANG[1,2], Kang ZHAO[3], Ming LIU[1*], Hanjun JIANG[3], Zhihua WANG[3], Zongwang ZHANG[4], Huili KAN[4], Ping CHEN[5] & Binjie ZHU[5]

[1]*State Key Laboratory on Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing* 100083*, China;*
[2]*University of Chinese Academy of Sciences, Beijing* 100049*, China;*
[3]*Tsinghua National Laboratory for Information Science and Technology, Institute of Microelectronics, Tsinghua University, Beijing* 100083*, China;*
[4]*Liaocheng People's Hospital, Liaocheng* 252000*, China;*
[5]*Beijing Yiemed Medical Technology Co., Ltd., Beijing* 100098*, China*

Received 20 May 2019/Accepted 2 August 2019/Published online 3 February 2021

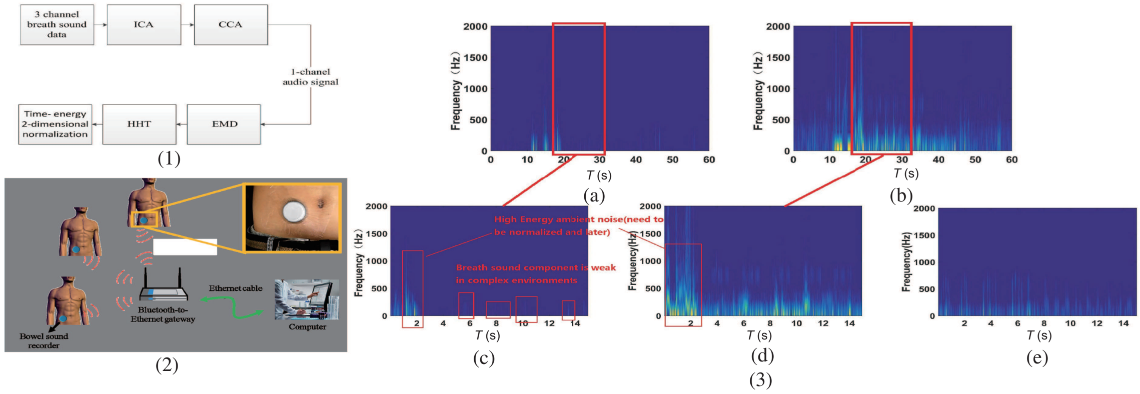
Citation Wang S K, Zhao K, Liu M, et al. Using breath sound data to detect intraoperative respiratory depression in non-intubated anesthesia. Sci China Inf Sci, 2021, 64(3): 134101, https://doi.org/10.1007/s11432-019-1510-6

Respiration is a vital function for every living thing. In clinical practice, any form of the respiratory dysfunction, if not detected and corrected in a timely manner, will have serious consequences. Conventional methods of monitoring the respiratory conditions of patients undergoing non-intubation anesthesia include stethoscope auscultation and peripheral capillary oxygen saturation SpO_2 monitoring. SpO_2 monitors the saturation of peripheral capillary hemoglobin, and its decrease indicates the occurrence of hemoglobin desaturation. However, hemoglobin desaturation has a time lag with respect to the occurrence of respiratory depression, and may occur later than respiratory depression [1]. In recent years, thoracic electrical bio-impedance (TEB) has been developed for respiratory monitoring in non-intubated anesthesia [2]. It has been shown to be effective, but when the impedance technique is used for respiratory monitoring, disturbances of the blood flow to the heart and the large blood vessels in the chest can be significantly disturbed. It has been shown that in specific populations, such as obese patients, the accuracy of the impedance method for the estimation of respiratory rate is lower than that based on acoustic sound estimation [3]. Breath sound auscultation is one of the best ways to monitor respiratory conditions. Because of the non-invasive nature of auscultation, and its low cost, it is an intuitive, valuable monitoring method to replace previously described methods for monitoring respiratory conditions during non-intubation anesthesia. Many research groups have studied the classification of normal and abnormal breath sounds, and some research groups have focused on the preprocessing of breath sound data, including the use of de-noising filters, multi-channel data fusion, and visualization of breath sound analysis results [4]. In general, however, the current research field of breath sound data processing is still immature, and the data used are rarely of clinical origin. The problem addressed in this study is the use of breath sounds to detect respiratory depression after anesthetic injection in non-intubation surgery.

Data collection and modeling. The data used in this research are collected during gastroscopic surgery. This surgery does not involve intubation, and is likely to cause respiratory depression. For details please see the video accompanying this study. The signal analysis of breath sound collection is a blind source separation problem. The basis of this research is the concept that the detection of respiratory depression must be related to a decrease in respiratory intensity. The problem is then naturally transformed into how to find the characteristics of breath sounds associated with breathing intensity. The most intuitive feature is the curve of the breath sound energy over time. Normalized energy characteristics can shield against the interference of human voices in the same frequency band as the breathing sound, to a certain extent. Our hypothesis is that the normalized breath sound energy is lower, and the time-energy curve is somewhat different from the time-energy curves of healthy people. If demonstrated, this should be a reliable biomarker.

Algorithm flow. Figure 1(2) shows a flow chart of the algorithm for breath sound feature extraction proposed in this study. There are three parts of to the signal: the breath sound signal obtained from the left back of the chest, the breath sound signal obtained from the right back, and ambient noise. In this study we applied a standard independent component analysis (ICA) with a Gaussian kernel function chosen for nonlinearity. The data, after passing through the ICA, has three segments, one of which is Gaussian noise, and the other two are matched with a template of normal breath sounds developed using all of the normal breath sound data

* Corresponding author (email: liuming@semi.ac.cn)



(4)

ACC/RECALL		Power spectrum	MFCC	This work
		Traditional features used		
Classification methods	KNN	0.32/0.3	0.52/0.1	0.44/0.2
	LR	0.6/0.5	0.56/0.5	0.64/0.5
	SVM-linear	0.68/0.4	0.6/0.4	0.52/0.5
	SVM-MLP	0.62/0.5	0.6/0.5	0.68/0.6

Figure 1 (Color online) (1) Algorithm flow. (2) Monitoring system architecture [7]. (3) Comparison results before and after ICA: (a) 1-minute raw signal segment power spectrum of the breath sound; (b) simultaneous segment selected after ICA processing; (c) zoom in on the selected part of (a); (d) simultaneous corresponding part of (b); (e) power spectrum of the normal breath sound template from [5]. (4) Algorithms comparison result.

in the ICBHI database [5]. After canonical correlation analysis (CCA) matching with the power spectrum, as shown in Figure 1(e), we select the most similar ICA component and apply empirical mode decomposition (EMD). We have prior knowledge that the required breath sound frequency band is 100–800 Hz [6], around 400 Hz. Therefore, we perform FFT on the result of each intrinsic mode function (IMF). If the obtained FFT value is less than 1.0 at 100 Hz, the components of the respiratory sound signal are almost no longer contained in the subsequent decomposition, so the EMD process is interrupted.

$$x(t)=\mathrm{Re}\sum_{i=1}^{n}a_i(t)\mathrm{e}^{\mathrm{j}\phi(t)}=\mathrm{Re}\sum_{i=1}^{n}a_i(t)\mathrm{e}^{\mathrm{j}\int w_i(t)\mathrm{d}t}, \tag{1}$$

where Re represents the imaginary part. Eq. (1) expresses the signal amplitude in a three-dimensional space as a function of time and instantaneous frequency, thereby enabling the extraction of the Hilbert amplitude spectrum of the signal, as in (2). The instantaneous energy density can be obtained by integrating the square of the amplitude of the Hilbert spectrum, allowing the calculation of the exact energy variation characteristic used in this study.

$$H(\omega,t)=\left\{\begin{array}{l}\mathrm{Re}\sum\limits_{i=1}^{n}a_i(t)\exp(\mathrm{j}\int\omega_i(t)\mathrm{d}t),\ \omega_i(t)=\omega\\ 0,\ \omega_i(t)\neq\omega\end{array}\right\}, \tag{2}$$

$$\mathrm{IE}(t)=\int_{\omega}H^2(\omega,t)\mathrm{d}\omega. \tag{3}$$

Breath sound time-energy dual-scale normalization. The intensity of a patient's breath sounds under anesthesia should decrease over time, and eventually increase after the drug effect wears off. However, each person's condition needs to be normalized, because each patient's physiological conditions are different. An indicator of respiratory depression for most people should be the moment that the respiratory rate is below a certain ratio, instead of the absolute value being below a certain value. The result of ICA processing is a linear amplification or reduction of the absolute value of the amplitude. The signal should also be normalized over time, because the intraoperative environment cannot be completely silent. The problem of non-Gaussian ambient noise needs to be solved, and the feature dimension can also be reduced, to facilitate subsequent classification. The normalization method used in this study is to convert the time-frequency Hilbert spectrum of a five-minute sampling rate of 4000 Hz into 150 points, with one point representing the energy average of two seconds. We estimate that breath sound energy accounts for 80%–90% of the entire energy of the signal. In this study we remove the first 10%–15% of the energy in the whole two seconds, and the corresponding time length is also removed to obtain a normalized energy value (we present an example in the supporting information). This feature extraction process can solve the problem of normalized energy characteristics caused by uneven distribution of noise interference in different time periods.

Results. Assistant physicians collected breath sound data from 25 patients undergoing gastroscopic surgery in Liaocheng people's hospital from May to August 2018. The type of surgery, and the operating room and microphone [7] used are the same. Each patient's breath sound characteristics consist of two parts: the normalized energy characteristic result of the ICA component processed by the algorithm described, and the result obtained by the same algorithm applied to the original respiratory sound data obtained from the left back. The two parts are combined into a 250 point characteristic curve, which is used in a subsequent SVM classification with leave-one-out validation. The SVM kernel function used is a multilayer perceptron with a default scale [1 −1]. The accuracy (ACC) of the experiment is 0.68, and the recall rate is 0.6. We find few reports of the use of intraoperative breath sound data to determine respiratory depression. We reproduce the results of features reported in other articles, and compare the results of different classification methods (Figure 1(3)). MFCC indicates Mel-Frequency Cepstral coefficients. The power spectrum method uses the industry wide method of shorttime Fourier transform (STFT) with a sample rate of 4000 Hz, frame length of 60 points, frame shift of 15 points, and FFT points of 256 points. The Fisher's linear discriminant (FLD) method is used to reduce the dimensions and then carry out subsequent classification. The classification methods selected are k-nearest neighbors (KNN), linear regression (LR), linear kernel SVM, and MLP kernel SVM. The judging criteria use two parameters, accuracy and recall. The feature extraction algorithm proposed in this study performs best using an SVM classification with an MLP kernel. The features used in this study produce better classification performance than traditional features.

Conclusion. In this study we designed a set of algorithms based on the energy characteristics of breath sounds, and used clinical gastroscopy data for verification. The algorithms have been proved to be effective, and the proposed energy-time curve feature produced better classification performance than traditional features.

Acknowledgements This work was supported by National Key Technologies R&D Program (Grant No. 2017YFB0405604), Key Research Program of Frontier Science, Chinese Academy of Sciences (Grant No. QYZDY-SSW-JSC004), Basic Research Project of Shanghai Science and Technology Commission (Grant No. 16JC1400101), and Beijing S&T Planning Task (Grant No. Z161100002616019).

Supporting information Videos and other supplemental documents. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Ahmed I S, Aziz E, Newton N. Connection of capnography sampling tube to an intravenous cannula. Anaesthesia, 2005, 60: 824–825

2 Frasca D, Geraud L, Charriere J M, et al. Comparison of acoustic and impedance methods with mask capnometry to assess respiration rate in obese patients recovering from general anaesthesia. Anaesthesia, 2015, 70: 26–31

3 Rizal A, Hidayat R, Nugroho H A, et al. Pulmonary crackle feature extraction using tsallis entropy for automatic lung sound classification. In: Proceedings of 2016 1st International Conference on Biomedical Engineering (IBIOMED), Yogyakarta, 2016

4 Palaniappan R, Sundaraj K, Sundaraj S. Artificial intelligence techniques used in respiratory sound analysis—a systematic review. Biomed Tech, 2014, 59: 7–18

5 Rocha B M, Filos D, Mendes L, et al. A respiratory sound database for the development of automated classification. In: Precision Medicine Powered by pHealth and Connected Health. Berlin: Springer, 2017. 33–37

6 Palaniappan R, Sundaraj K, Sundaraj S. Artificial intelligence techniques used in respiratory sound analysis–a systematic review. Biomed Tech, 2014, 59: 7–18

7 Yin Y, Jiang H J, Feng S L, et al. Bowel sound recognition using SVM classification in a wearable health monitoring system. Sci China Inf Sci, 2018, 61: 084301