2022 44th Annual International Conference of
the IEEE Engineering in Medicine & Biology Society (EMBC)
Scottish Event Campus, Glasgow, UK, July 11-15, 2022

A Wearable Swallowing Recognition System Based on Motion and Dual Photoplethysmography Sensing of Laryngeal Movements

Ying Zhang, Huaiyu Zhu, Haipeng Liu, Dingchang Zheng, Shaomin Zhang, and Yun Pan*

Abstract—**Swallowing recognition is the leading step in the evaluation of dysphagia which seriously affects people's life. Current medical swallowing monitoring methods require an in-hospital environment and overly rely on professional knowledge of the medical staff. In this study, we developed a wearable swallowing recognition system that consists of an on-neck wearable swallowing sensing device and a data processing module on a host computer. The wearable device collects inertial signals including acceleration and angular velocity, as well as dual photoplethysmography (PPG) signals based on infrared and green light from the neck. A novel processing framework for dual PPG signals is proposed to extract and enhance the laryngeal motion component introduced by swallowing activities in the data processing module. The laryngeal motion component of dual PPG signals together with the preprocessed inertial signals are further used for feature extraction to proceed swallowing recognition based on random forest classifier. We collected data from 32 healthy subjects in the center and side positions on the neck using our system to analyze their swallowing activities. As a result, we achieved a high average area under curve (AUC) of the swallowing recognition by 86.6%. We also find the sensing position has a significant impact on gender-specific swallowing recognition performance, as the center position was better for females (92.9%), while the side position was better for males (87.6%). The results indicate that the proposed system could achieve high integrity and good performance, which is helpful for the future swallowing research.**

I. INTRODUCTION

Swallowing is a common daily activity, which involves the movement of multiple cartilage and tissues, especially thyroid cartilage and cricoid cartilage. In clinical practice, the videofluoroscopic / videoendoscopic swallowing study (VFSS / VESS) are commonly used to monitor subjects' swallowing. However, such methods are limited by in-hospital environment and cannot fully reflect subjects' swallowing state in daily life. Therefore, it is of great significance to study swallowing monitoring system with good user experience and the ability of daily monitoring, where mobile health technology could meet the requirements.

Nicholls et al. [1] designed a skin like flexible polymer sensor for surface electromyography acquisition and a game feedback interface on computer to realize the functions of swallowing monitoring and swallowing behavior correction. Kuramoto et al. [2] utilized microphones and inertial sensors to detect changes in neck angle during swallowing with a wireless visual feedback device to convey the swallowing state. Nagae et al. [3] implemented a swallowing monitoring system to monitor real-time swallowing behavior in daily life and distinguish swallowing from speaking and coughing. Wei et al. [4] considered that the movement of laryngeal cartilage introduced "noise" into laryngeal PPG signal, and they used the "noise" and inertial signals to identify swallowing behavior.

In this study, we developed a wearable swallowing recognition system based on dual photoplethysmograghy (PPG) and inertial signals sensing from laryngeal movements that includes a novel dual PPG motion component extraction framework, and analyzed the swallowing behavior using the laryngeal motion information extracted from these signals. A total 470 10-second laryngeal movement data from 32 healthy subjects (14 females, 18 males) in the center and side positions were collected for system training and validation. The center position is located in the depression between thyroid cartilage and cricoid cartilage [5], and the side position is with the dense distribution of cricothyroid artery. The system achieved a high average area under curve (AUC) of the swallowing recognition by 86.6%. Meanwhile, the results indicates that the sensing position has a significant impact on the performance of gender-specific swallowing recognition, as the center position is better for the female swallowing recognition, while the side position is better for the male one.

II. SYSTEM DEVELOPMENT

We designed and implemented a wearable swallowing recognition system, as shown in Fig. 1. The inertial sensing data including 3-axis acceleration and 3-axis angular velocity, as well as the dual PPG data during swallowing and at rest were collected by the wearable device. These data were transmitted wirelessly to a host computer for data processing. The original inertial signals were first preprocessed by a 12-level "sym8" wavelet decomposition and reconstructed without the approximation coefficient to eliminate the baseline drift. And the signals underwent a 10-point moving window mean smoothing to eliminate the burr. We took the modulus of acceleration and angular velocity respectively to reduce the feature dimension. On the other hand, the dual PPG signals

*This research was supported in part by Zhejiang Provincial Natural Science Foundation of China under Grant No. LQ21F010016 and in part by Zhejiang Provincial Key Research and Development Program of China (2021C03027). (Corresponding author: Yun Pan.)

Ying Zhang, Huaiyu Zhu, and Yun Pan are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: {zhangy21, zhuhuaiyu, panyun} @zju.edu.cn).

Haipeng Liu and Dingchang Zheng are with Research Centre of Intelligent Healthcare, Coventry University, Coventry, United Kingdom (e-mail: {ad4828, dingchang.zheng} @coventry.ac.uk).

Shaomin Zhang is with Qiushi Academy for Advanced Studies, Zhejiang Provincial Key Laboratory of Cardio-Cerebral Vascular Detection Technology and Medicinal Effectiveness Appraisal, Zhejiang University, Hangzhou 310027, China (e-mail: shaomin@zju.edu.cn).

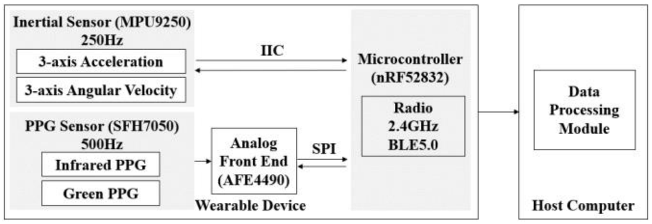


Figure 1. Configuration of the swallowing recognition system

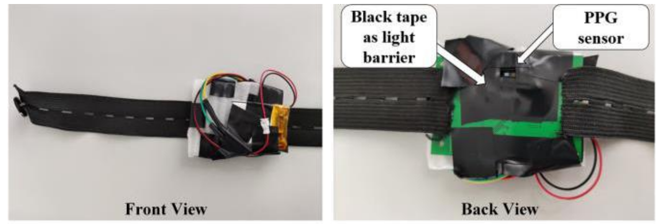


Figure 2. The wearable device for swallowing signals acquisition

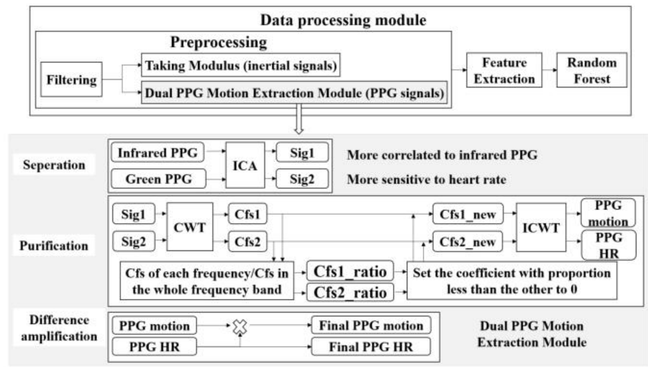


Figure 3. Dual PPG motion component extraction configuration

were filtered by band-pass filter with a pass-band between 0.5 Hz and 4 Hz, and the swallowing motion components were extracted by independent component analysis (ICA), time-frequency proportion subtraction, and time-domain amplitude multiplication. The swallowing segments and resting segments were identified and distinguished by random forest.

A. Wearable Device Design

The on-neck wearable laryngeal movement sensing device is shown in Fig. 2. The device could acquire swallowing activities via inertial and dual PPG signals and transmit them to host computers by Bluetooth. Two pieces of elastic bands were integrated with the printed circuit board to achieve neck wearing function and improve the portability. The black tapes were applied to isolate the interference of ambient light.

B. Dual PPG Motion Component Extraction Module

Previous studies showed that PPG with long wavelength such as infrared light can be disturbed by the movement of deep tissues due to its stronger ability of penetration [6], which brings more motion component. And infrared light can reach deeper skin tissue than short light such as green and blue light [7]. Therefore, Zhang et al. [8] took the infrared PPG as the micro motion reference, and the green PPG with greater signal motion artifact ratio as the main signal to extract the heart rate (HR). The methods of normalized time-frequency subtraction after continue wavelet transform (CWT) and the time-domain decomposition and reconstruction were adopted to eliminate the motion noise and reconstruct the heart rate component. Lee et al. [9] adopted ICA to decompose the HR components from the PPG signals of 12 channels (4 positions with infrared, red, and green light), and then applied truncated singular value decomposition (SVD) to the further motion artifacts reduction. Additionally, in [10], the inertial signals and green PPG served as motion reference, and the red PPG was reconstructed to remove motion artifacts from based on wavelet method.

Inspired by these studies, on the contrary, a framework was proposed in this study to extract motion components from PPG signals with different wavelengths for swallowing recognition. In particular, the motion component extraction module of this study mainly includes three steps as shown in Fig. 3. We first carried out ICA to preliminarily decompose the motion components and HR components in the infrared and green light PPG. Thus, we obtained two signals which are with greater correlation with the original infrared light PPG signal, and more sensitive to heart rate respectively. Then, to further reduce the coupling of the two signals, based on the difference between them in time-frequency domain, the wavelet coefficients proportion the two signals in the whole frequency band were compared with the other to retain the wavelet coefficients with greater proportion, and then the two signals described as "PPG motion" and "PPG HR" in Fig. 3 were reconstructed respectively. Finally, since the swallowing and resting segments of the motion component ("PPG motion" in Fig. 3) both had high fluctuation, it was multiplied by the signal of the HR component ("PPG HR" in Fig. 3) with small overall fluctuation to enlarge the gap between swallowing and resting segments. As a result, the final motion component ("Final PPG motion" in Fig. 3) was extracted.

C. Feature Extraction and Model Training

In this system, we collected signals for 10 seconds at a time, while the basic swallowing recognition unit was set to 1 second according to the generally approximate duration of single healthy swallowing with 5-mL bolus of 1 second [11]. We extracted 32 features from each of the signals including the modulus of three-axis acceleration, the modulus of three-axis angular velocity and the PPG motion component in time domain, frequency domain, wavelet domain and information domain. In order to reduce individual differences, we calculated the ratio of features of each 1-second short segment in the 10-second sample, so as to focus only on the relative fluctuation per second in 10 seconds. The random forest screened out the features below a certain threshold according to the out-of-bag error (OOBE). Among the final features, the 10 most commonly used features were the proportion of the following variables: the range of acceleration, the range of angular velocity, the range of PPG motion component, the mean value of angular velocity, the mean value of PPG motion component, the interquartile range of angular velocity, the interquartile range of PPG motion component, the standard deviation of acceleration, the standard deviation of angular velocity, and the standard deviation of PPG motion component.

We applied 5-fold cross validation in the intra-group training and testing of random forest. In each fold, the best hyperparameter combination was selected according to F1-score through grid search on the training set for each 10-second sample. The inputs of random forest were the features of each 1-second segment to classify whether it contained swallowing or not (1 bit "1" or "0"). Finally, a 10 bit "0" and "1" sequence was obtained for 10s samples.

III. Experiments

A. Experiment Setup

In the experiment, the device was used to collect inertial signals and dual PPG signals for swallowing detection in two positions as described in Fig. 4 (a). One of the positions is the median depression between the thyroid and cricoid cartilage, i.e., the center position, and the other is the intersection of the horizontal line of thyroid cartilage and the circular thyroid artery, i.e., the side position. The former position with more bony structures reveals more swallowing information for inertial sensors, while the latter reflects stronger PPG signals due to its more abundant artery distribution. A total of 32 healthy subjects participated in the experiment, including 14 females and 18 males with an age of 22.6±2.1. This study was approved by the Medical Ethics Committee of the Medical College of Zhejiang University (Document No. 2020-107).

B. Data Overview

Subjects were asked to sit upright with the device fixed by an elastic band around their necks and the PPG sensor fitted the center or side position. For each position, the subjects swallowed a mouthful of 5 mL water at any time within 10 seconds after the serial port indicator on the development board controlled by a button lighted on. A single measurement lasted for 10 seconds. The above steps were repeated 10 times, and the subjects remained stationary in another two repeated measurements. The system in measurement is shown in Fig. 4 (b). The data were divided into four groups according to four conditions with the combinations of positions and gender as shown in Table I. Each 10-second sample was labeled manually with the minimum unit of labeling set to 1 second to form a 10-bit sequence, and only when each second in this sample contained no swallowing, could it be identified as a negative sample. Otherwise, it was a positive sample.

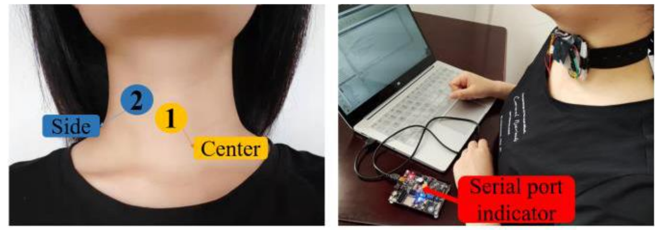
IV. Results and Discussion

A. Results

In this study, all metrics were calculated based on the recognition results listed in Table II [4] except the metrics p-Accuracy and n-Accuracy. These two metrics are defined as the proportions of TP in all 10-second samples containing swallowing and the proportions of TN in all 10-second resting samples respectively. The swallowing recognition results of male and female in the two positions are shown in Table III. Overall, the swallowing recognition performance of the system in female group with an average AUC of 91.0% (see the italic and bold in Table III) is better than that in male with an average AUC of 82.2%. At the same time, for female, the center position is a more suitable position to collect swallowing signals due to its better recognition performance with an AUC of 92.9%. While the side position is better for male with an AUC of 87.6%.

B. Feature Analysis and Discussion

In order to explain the better data measurement position for each gender, we calculated the OOBE of each feature in the random forest model for four groups (the higher the OOBE, the greater the importance of the feature), and listed five features with the highest OOBE, as shown in Table IV.



(a) (b)

Figure 4. (a) Data collected positions, (b) System in measurement

TABLE I. Distribution of Dataset

Groups	Center Position				Side Position			
	Female		Male		Female		Male	
Samples (10s)	P[a]	N[a]	P	N	P	N	P	N
Number	88	13	117	17	84	17	114	20
Sum	101		134		101		134	

a. P and N refer to positive and negative samples collected during swallowing and at rest respectively

TABLE II. Definition of Recognition Results

Results	Definition
TP	Number of positive samples with each 1-sec segment correctly recognized
TN	Number of negative samples correctly recognized
FP	Number of negative samples incorrectly recognized + Number of positive samples with misjudged swallowing duration
FN	Number of positive samples identified as negative

The difference of swallowing recognition performance under four experimental groups are analyzed. For the female subjects, the performance with an AUC of 92.9% in the center position is better, which may be because the center position is in the main area of swallowing and is more affected by swallowing. The cartilage movement caused by swallowing is easier to be collected by the inertial sensor in the center position with more abundant cartilage and fewer arteries, which makes the inertial signal here better reflect swallowing. Thus, the five most important features of female in the center position in the Table IV are mainly inertial features with a ratio of 4/5. In the side position, although the influence of swallowing reflected by all signals is weak, the arteries increase and the cartilage decreases, which is beneficial for PPG signals reflecting swallowing. As a result, for female group, the significance of PPG features increases to compensate for the weakening of swallowing effects in the side position as the number of PPG features in the top 5 important features increases from 1 in the center position to 3 in the side position.

The performance on female subjects is better for two reasons. Firstly, the poor fit between the board and the neck caused by the protrusion of male's thyroid cartilage not only weakens the sensitivity of inertial signal to swallowing, but also leads to the decline of PPG signal quality as a result of light leakage. For the male subjects, the worse performance in the center position may be the result of the poorer fit of the board and the lower PPG energy.

TABLE III. EVALUATION ON GENDER AND POSITION

Groups		**Evaluation Metrics**						
Gender	**Position**	**TPR**	**TNR**	**Accuracy**	**F-measure**	**AUC**	**p-Accuracy**	**n-Accuracy**
Female	Center	100.0%±0.0%	70.0%±27.4%	92.0%±9.1%	94.7%±6.4%	**92.9%±9.6%**	90.4%±11.2%	100.0%±0.0%
	Side	100.0%±0.0%	70.7%±34.9%	92.1%±9.1%	95.1%±5.5%	89.0%±16.0%	91.0%±9.9%	100.0%±0.0%
Average for female		100.0%±0.0%	70.3%±31.2%	92.0%±9.1%	94.9%±5.9%	**91.0%±12.8%**	90.7%±10.6%	100.0%±0.0%
Male	Center	100.0%±0.0%	47.2%±27.6%	88.9%±6.9%	93.3%±4.0%	76.9%±11.5%	89.1%±7.2%	83.3%±23.6%
	Side	100.0%±0.0%	54.9%±15.3%	88.8%±4.5%	92.9%±3.1%	**87.6%±9.3%**	86.8%±5.3%	100.0%±0.0%
Average for male		100.0%±0.0%	51.1%±21.3%	88.8%±6.1%	93.1%±4.0%	82.2%±10.6%	88.0%±7.0%	91.7%±2.4%
Average		100.0%±0.0%	60.7%±26.3%	90.4%±7.4%	94.0%±4.7%	86.6%±11.6%	89.3%±8.4%	95.8%±5.9%

TABLE IV. THE FIVE MOST IMPORTANT FEATURES OF EACH GROUP

Groups							
Female-Center		**Female-Side**		**Male-Center**		**Male-Side**	
Feature Name	**OOBE**[b]	**Feature Name**	**OOBE**	**Feature Name**	**OOBE**	**Feature Name**	**OOBE**
Angular velocity (ω) mean	0.7651	ω range	0.6712	ω mean	0.7917	ω mean	0.6551
ω standard deviation (std)	0.6527	PPG_m std	0.6071	Acceleration range	0.7577	PPG_m std	0.5971
PPG motion (PPG_m) range	0.5795	PPG_m range	0.6065	PPG_m std	0.7049	ω iqr	0.5074
ω range	0.5352	ω mean	0.5778	PPG_m range	0.6270	ω std	0.5027
ω interquartile range (iqr)	0.3726	PPG_m mean	0.5521	ω std	0.6035	PPG_m mean	0.4336

b. OOBE refers to out-of-bag error

Different from the situation that the model in female group tends to use inertial signals in the center position and PPG signal in the side position, which is the natural result of physiological structure, the ratios of inertial features and PPG features in the two positions of male in the Table IV are both 3:2, which indicates that the model on the male dataset "has to" take a compromise feature selection measure as a result of the poor signals. Secondly, the arteries of the female group may be more superficial [12], leading to stronger PPG that can be collected more under the influence of swallowing.

V. CONCLUSION

We developed a wearable swallowing recognition system based on inertial signals and dual PPG signals collected on the neck that reflect laryngeal movement during swallowing. In the data processing module, a novel processing framework for dual PPG signals is proposed to extract and enhance the motion in swallowing. The system was used to collect and process data from 32 healthy subjects in the center and side position on the neck which achieved a high average AUC of the swallowing recognition of 86.6%. In addition, we found that under our system, the better swallowing data acquisition positions related to performance are different for female and male. The center position is better for female with an AUC of 92.9%, while the side position is better for male with an AUC of 87.6%. The system is expected to be used for daily swallowing monitoring and research on swallowing disorders. There are still some limitations. For example, the low fit of the hard printed circuit board and neck leads to the deterioration of signal quality, and the feature selection and recognition algorithm is also relatively rough, which will be improved in the follow-up study.

REFERENCES

[1] B. Nicholls, C. S. Ang, C. Efstratiou, Y. Lee and W. -H. Yeo, "Swallowing detection for game control: using skin-like electronics to support people with dysphagia," in *IEEE Int. Conf. Pervasive Comput. and Commun. Workshops,* Hawaii, 2017, pp. 413-418.

[2] N. Kuramoto, D. Jayatilake, K. Hidaka and K. Suzuki, "Smartphone-based swallowing monitoring and feedback device for mealtime assistance in nursing homes," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Orlando, 2016, pp. 5781-5784.

[3] M. Nagae and K. Suzuki, "A neck mounted interface for sensing the swallowing activity based on swallowing sound," in *Proc. 33rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Boston, 2011, pp. 5224-5227.

[4] L. Wei, Y. Pan, Y. Zhong and R. Huan, "A Novel Approach for Swallow Detection by Fusing Throat Acceleration and PPG Signal," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Honolulu, 2018, pp. 4293-4296.

[5] Y. Zhong, Y. Pan, L. Zhang and K. -T. Cheng, "A wearable signal acquisition system for physiological signs including throat PPG," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Orlando, 2016, pp. 603-606.

[6] V. Vizbara, A. Sološenko, D. Stankevičius, V. Marozas, "Comparison of green, blue and infrared light in wrist and forehead photoplethysmography," in *Proc. 20th Int Conf. Biol. Eng.*, Kaunas, Nov 2016, vol. 17, no.1, pp.24-25.

[7] S. Han, D. Roh, J. Park and H. Shin, "Design of Multi-Wavelength Optical Sensor Module for Depth-Dependent Photoplethysmography," *Sensors*, vol. 19, no. 24, pp. 5441, Dec. 2019.

[8] Y. Zhang, S. Song, R. Vullings, D. Biswas, N. Simões-Capela, N. van Helleputte, C. van Hoof, and W. Groenendaal, "Motion Artifact Reduction for Wrist-Worn Photoplethysmograph Sensors Based on Different Wavelengths," *Sensors*, vol. 19, no. 3, pp. 673, Feb. 2019.

[9] J. Lee, M. Kim, H. -K. Park, and I. Y. Kim, "Motion artifact reduction in wearable photoplethysmography based on multi-channel sensors with multiple wavelengths," *Sensors*, vol. 20, no.5, pp. 1493, Mar. 2020.

[10] A. -M. Tăuțan, A. Young, E. Wentink and F. Wieringa, "Characterization and reduction of motion artifacts in photoplethysmographic signals from a wrist-worn device," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Milan, 2015, pp. 6146-6149.

[11] M. R. Hoffman, M. R. Ciucci, J. D. Mielens, J. J Jiang, and T. M. McCulloch, "Pharyngeal swallow adaptations to bolus volume measured with high-resolution manometry," *Laryngoscope*, vol.120, no.12, pp. 2367-2373, Dec. 2010.

[12] S. Seidenari, A. Pagnoni, A. Di Nardo, A. Giannetti, "Echographic evaluation with image analysis of normal skin: Variations according to age and sex," *Skin Pharmacol.*, vol. 7, no. 4, pp. 201-209, 1994.