

Hearables: Heart Rate Variability from Ear Electrocardiogram and Ear Photoplethysmogram (Ear-ECG and Ear-PPG)

Haozhe Tian, Edoardo Occhipinti, Amir Nassibi, and Danilo P. Mandic*

Abstract—This work aims to classify physiological states using heart rate variability (HRV) features extracted from electrocardiograms recorded in the ears (ear-ECG). The physiological states considered in this work are: (a) normal breathing, (b) controlled slow breathing, and (c) mental exercises. Since both (b) and (c) cause higher variance in heartbeat intervals, breathing-related features (SpO_2 and mean breathing interval) from the ear Photoplethysmogram (ear-PPG) are used to facilitate classification. This work: 1) proposes a scheme that, after initialization, automatically extracts R-peaks from low signal-to-noise ratio ear-ECG; 2) verifies the feasibility of extracting meaningful HRV features from ear-ECG; 3) quantitatively compares several ear-ECG sites; and 4) discusses the benefits of combining ear-ECG and ear-PPG features.

I. INTRODUCTION

Non-invasive Heart Rate Variability (HRV) monitoring relies on the Electrocardiogram (ECG) or Photoplethysmogram (PPG) recordings. Wearable devices, which do not interfere with the everyday activities of subjects, are ideal for the daily monitoring of HRV. Among wearable devices, ear-based devices, also known as hearables, are attracting increasing attention as they can be fitted to headphones or earplugs [1], [2], [3]. Also, given that the head is in a much more stable position with respect to the vital organs, compared to e.g. the wrist, signals recorded from the ears are less affected by artefacts and are therefore better suited for 24/7 monitoring. Recent work also proved the possibility of ear electroencephalogram artefact-removal [4], [5].

In [1], [2], the theoretical feasibility of measuring ECG from around the ear was proven. For real-life recordings, several ECG sites around the ear were quantitatively compared in [6], and the ECG measured from the mastoid showed adequate quality for HRV monitoring. Besides ear-ECG, recent research also proved the benefits of ear-PPG against finger PPG: 1) it shows stronger breathing variation and is, therefore, more suited for extracting breathing-related features [7], [8]; and 2) it has a significantly faster response time to drops in SpO_2 [3].

*Edoardo Occhipinti was supported by UKRI Centre for Doctoral Training in AI for Healthcare grant number EP/S023283/1

Haozhe Tian, Amir Nassibi and Danilo P. Mandic are with the Department of Electrical and Electronic Engineering, Edoardo Occhipinti is with the UKRI Centre for Doctoral Training in AI for Healthcare, Imperial College London, London SW7 2AZ, UK (email: {haozhe.tian21, edoardo.occhipinti16, a.nassibi15, d.mandic}@imperial.ac.uk).

H. Tian, E. Occhipinti, A. Nassibi and D. P. Mandic, "Hearables: Heart Rate Variability from Ear Electrocardiogram and Ear Photoplethysmogram (Ear-ECG and Ear-PPG)," Proceedings of the 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2023, pp. 1-5, doi: 10.1109/EMBC40787.2023.10340371.

Despite the proven possibility of ECG measurement from around the ear, the feasibility of ear-HRV monitoring has not yet been verified on real-life data. The ear-ECG feasibility studies have demonstrated the inherent advantages of ear-ECG and ear-PPG configurations, but have also highlighted that the applications face two challenges. First, many physiological states are difficult to distinguish by HRV alone. For instance, both controlled slow breathing at 0.1Hz frequency and mental exercise cause increased HRV [9], [10]. The scenario where both states are present was not widely discussed. Second, due to lower signal amplitudes, ear-ECG and ear-PPG have relatively low signal-to-noise ratios (SNR). The matched filter and Hilbert transform (MFHT) algorithm [11] performed well on the detection of R-peaks in low-SNR wearable ECG signals, though requiring a manually selected QRS pattern to act as a template.

This study contributes to the topic of ear physiological state monitoring in four aspects:

- It uses higher-SNR arm ECG to pre-train MFHT template and quantitatively compares the method to the classic Pan-Tompkins' method (PT) [12].
- To the best of our knowledge, this is the first time to verify the feasibility of ear-ECG HRV monitoring.
- It includes both controlled slow breathing and mental exercises in the protocol to simulate real-life scenarios.
- It proves the benefits of using ear-PPG to aid physiological state classification.

II. CONFIGURATION

A. Recording Set-up

For this study, 7 healthy subjects (5 males and 2 females) aged 20-30 were recruited. Multiple ear-ECG sites and one

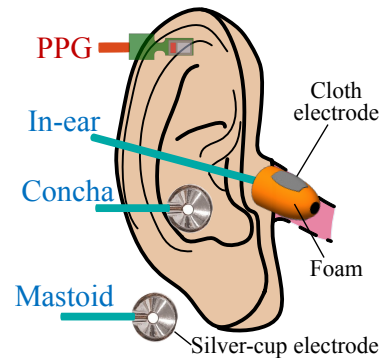


Fig. 1. Sensor placement used for recording ear-ECG and ear-PPG signals. The in-ear ECG signal was recorded via a flexible cloth electrode attached to a viscoelastic earplug.

ear-PPG site were recorded. The three ear-ECG sites, which enabled site-specific signal comparison, were: in-ear, concha, and mastoid (marked in blue in Fig. 1). For all ear-ECG sites, the same site at the other ear was used as the respective reference. Prior to starting the recording, the mastoid and the concha were abraded by an abrasive electrode gel, and silver cup electrodes were held in position with conductive paste fixed with medical tape. For in-ear ECG, viscoelastic earplugs and cloth electrodes [13] were used. The cloth electrodes were covered with conductive gel before inserting to ensure closer fit and lower impedance. The PPG sensor (marked in red in Fig. 1) was taped behind the right ear. All signals were recorded simultaneously via our custom-made bio-amplifier, the BioBoard.

B. Protocol

As mentioned in Section I, this work collected responses to both controlled slow breathing and mental exercises in the recording protocol to simulate real-life scenarios. The recordings were guided by specially designed MATLAB apps. The use of interactive apps ensured an accurate partition of recording sessions. Subjects were also free to progress at their will, which reduced the outlier data caused by subjects not being familiar with the procedures.

The recording protocol contained three sections as shown in Fig. 2. Before starting the recording, subjects rested for 10 minutes. In session (a), normal breathing, subjects sat breathing normally for 5 minutes. In session (b), slow breathing, subjects breathed with a fixed breathing rate (0.1Hz frequency) following an on-screen indicator. After sessions (a) and (b), subjects rested for 10 minutes. Before proceeding to session (c), mental exercises, subjects were allowed free practice of the interface. When session (c) began, subjects were told to accurately answer as many mental exercises as possible in 5 minutes. Two kinds of mental exercises were given: colour-word test [14] and Raven's standard test [15]. Before starting the protocol, each subject gave written consent. The protocol has been approved under the IC ethics committee approval JRCO 20IC6414.

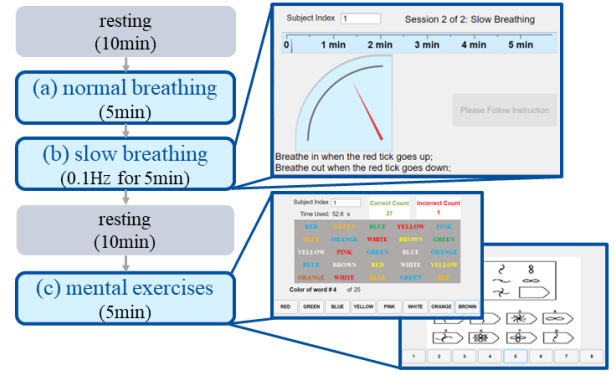


Fig. 2. The recording protocol used in this research. The whole process was guided by an interactive app to ensure consistency among all subjects.

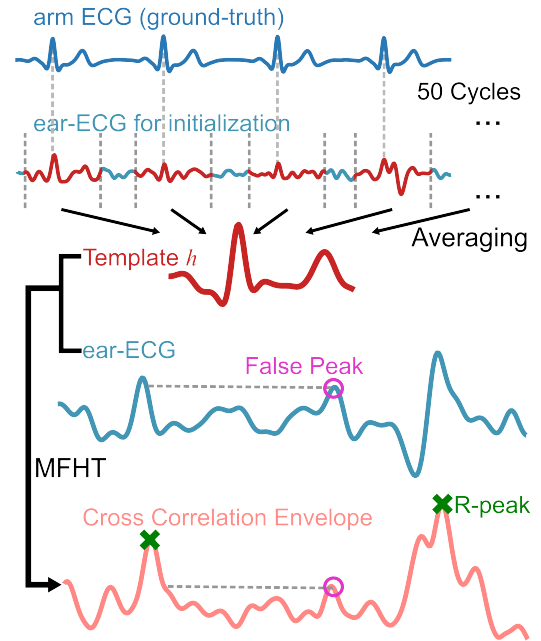


Fig. 3. Illustration of the mMFHT method [11] for R-peak extraction from ear-ECG. An improvement in amplitude difference between the ear-ECG R-peak and the false peak was observed.

TABLE I
EXTRACTED FEATURES

ECG Time Domain (t)	
HR	heart rate per minute
RMSSD	root mean square of adjacent RR interval difference
pNN50	ratio of RR intervals with >50ms adjacent difference
SDNN	standard deviation of RR intervals
ECG Frequency Domain (f)	
LF	power of the lower frequency band (0.04 to 0.15Hz) of RR interval signal
HF	power of the higher frequency band (0.15 to 0.4Hz) of RR interval signal
LF/HF	ratio between LF and HF
PPG (p)	
SpO ₂	peripheral capillary oxygen saturation
MBI	mean breathing interval

III. METHODOLOGY

A. ECG Features

As summarized in [10], [16], heart rate (HR) and HRV features proved effective in distinguishing mental exercises from normal states. To calculate HRV features, the R-peaks of ear-ECG should be extracted, while the RR interval signal is used to calculate the features in Table I.

Before starting, all ECG signals were pre-processed using band-pass filters between 2 and 12Hz. Different from [11], zero-phase filters were used in this work to minimize phase distortion. The multi-channel MFHT (mMFHT) was used to locate the R-peaks. First, mMFHT identified 50 R-peaks (using MATLAB `findpeaks` function) in high SNR arm ECG reference. Let the index of the i -th arm ECG R-peak

be $n_i, i \in 0, 1, \dots, 49$. Since the ear-ECG signal, x , and arm ECG were measured synchronously, the i -th ear-ECG R-peak also has an index n_i . Under a 200Hz sampling rate, the ear-ECG template used for matched filtering was

$$h = \frac{1}{50} \sum_{i=0}^{49} x[n_i - 69 : n_i + 50] \quad (1)$$

The template h had a length of 0.6 seconds, which was sufficient to cover the QRS pattern. After the determination of h , the arm ECG reference was no longer used. The subsequent procedures of mMFHT were similar to MFHT in [11]. The cross-correlation between the determined QRS template h and ear-ECG signal x was calculated as

$$y[n] = \sum_k h[k]x[n+k] \quad (2)$$

The Hilbert transform, y_h , of y was used to form a positive envelope signal, a , which was calculated as

$$a[n] = \sqrt{y^2[n] + y_h^2[n]} \quad (3)$$

The R-peaks of ear-ECG were extracted from signal a . As shown in Fig. 3, the mMFHT greatly improved the prominence of R-peaks against false peaks caused by noise. Based on R-peaks extracted by mMFHT, the time and frequency domain ECG features in Table I were calculated.

B. PPG Features

Since both session (b), slow breathing, and session (c), mental exercises, in the recording protocol (Fig. 2) cause more variance in RR intervals, we explored the possibility of using breathing-related features, SpO_2 and mean breathing interval (MBI), to aid the classification of the two. The calculation of SpO_2 and MBI is commonly based on PPG. Several works show that compared to finger PPG, ear-PPG has stronger breathing-related fluctuations and responds

faster to oxygen level changes [3], [7], [8]. The reason might be that the in-ear site is less affected by vasoconstriction and has a relatively constant distance from vital organs. This work extracts SpO_2 and MBI from ear-PPG.

The ear-PPG sensor (MAX30101 by Maxim Integrated) has both red (660nm) and infrared (880nm) light sources. Using the manufacturer suggested calibration [17], the SpO_2 was calculated as

$$\text{SpO}_2 = 104 - 17 \times \frac{\text{AC}_{\text{red}}/\text{DC}_{\text{red}}}{\text{AC}_{\text{infrared}}/\text{DC}_{\text{infrared}}} \quad (4)$$

The extraction of the AC amplitudes involved three steps. First, the signal was band-passed between 1 and 25Hz using a zero-phase IIR filter with a minimum degree that achieved 60dB stop-band attenuation. Second, the peaks and troughs of the filtered signal were identified using the MATLAB function `findpeaks`. To identify troughs, the signal was inverted. Two three-point median filters were applied to the extracted peaks and troughs to remove the artefact peaks. Third, the peaks and troughs were interpolated with splines using the MATLAB `interp1` function. A schematic of the interpolated peak signal and trough signal is shown in Fig. 4 (a). The interpolated peak signal subtracted by the interpolated trough signal gave the AC signal, which was averaged to give the AC amplitude. To extract DC amplitudes from the red and infrared signals, ear-PPG was low-pass filtered at 0.02Hz. The average of the filtered signal was the extracted DC amplitude.

This work extracted MBI by observing the baseline wanders of PPG signal [18]. As illustrated in Fig. 4 (b), the PPG baseline shows fluctuations corresponding to inhaling and exhaling. To obtain the PPG baseline, the raw ear-PPG signal was first band-pass filtered between 0.05Hz and 30Hz. Afterwards, a 3 seconds moving average filter was applied. Next, baseline peaks were found based on the premise that the minimum breathing interval was larger than 1 second, and then MBI was the average of PPG baseline peak intervals.

IV. RESULTS AND DISCUSSION

A. mMFHT vs. PT

The ear-ECG R-peak extraction method (mMFHT) proposed in Section III-A was compared with the standard PT algorithm [12]. Since the focus is ECG R-peaks instead of morphology, the arm ECG has adequate accuracy to serve as the ground truth. To ensure the validity of the evaluation, the arm ECG R-peaks were manually scrutinized. If an identified ear-ECG R-peak was within the 30ms range of a ground truth peak, the identification was considered correct. To illustrate the performance gap between mMFHT and PT, the ear-ECG with the lowest SNR (detail in Section IV-B) was used for comparison.

The metrics used for comparison were: *precision*, *recall*, F_1 , and mean offset. The *precision* and *recall* are calculated as

$$\text{precision} = \frac{TP}{TP + FP} \quad \text{recall} = \frac{TP}{TP + FN} \quad (5)$$

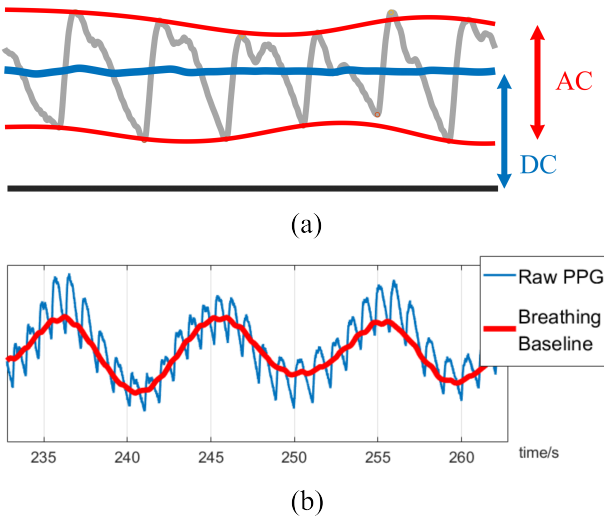


Fig. 4. Ear-PPG signals. (a) The AC and DC amplitudes of the ear-PPG signal. (b) Extracting MBI using the fluctuating ear-PPG baseline.

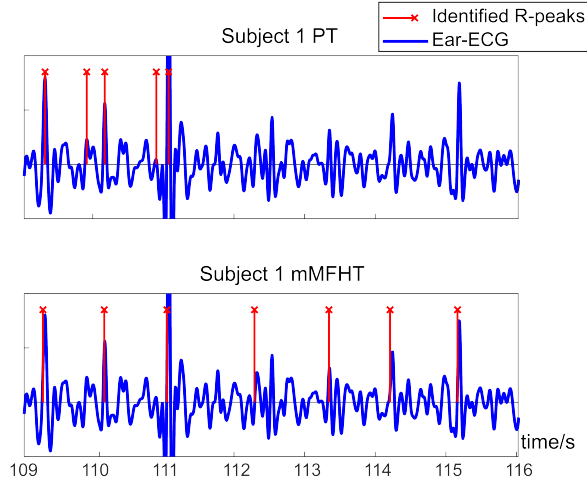


Fig. 5. Identification of R-peaks for subjects #1, #5, #6, and #7. Large artefacts caused errors in the adaptive threshold of the PT algorithm.

TABLE II
PERFORMANCE OF MMFHT [11] AND PT [12] ON IN-EAR ECG

Method	<i>precision</i>	<i>recall</i>	F_1	mean offset
PT	0.49	0.88	0.63	13ms
mMFHT	0.85	0.83	0.84	8ms

where TP is the number of correctly identified R-peaks, FP is the number of falsely identified R-peaks, and FN is the number of missed R-peaks. High *precision* and low *recall* means the method identifies fewer R-peaks but with higher confidence, and vice versa. The F_1 score combines precision and recall and is calculated as

$$F_1 = 2 * \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (6)$$

For subjects #1, #5, #6, and #7, mMFHT was able to identify R-peaks with average $F_1 = 0.58$. However, large artefacts caused errors in the adaptive threshold of PT, therefore PT was not able to identify any R-peak for large fractions of the recordings (Fig. 5). For subjects #2, #3, and #4, the average *precision*, *recall*, F_1 , and mean offset are shown in Table II. It was observed that, despite slightly lower *recall* (5%), mMFHT outperformed PT in the low-SNR setting.

B. Ear-ECG Site Comparison

Table III shows the comparison of different ear-ECG sites with arm ECG as the ground truth. The results in Table III shows the average over all 7 subjects, with the variance given in the brackets. In terms of R-peak extraction accuracy, it was observed that mastoid > concha > in-ear. The mastoid ECG had $F_1 > 0.95$, with a small cross-subject variance (0.03). The in-ear ECG, on the other hand, showed a big cross-subject variance. For subjects #2, #3, and #4, the F_1 of in-ear ECG ranged from 0.82 to 0.89, while for other subjects' F_1 ranged from 0.44 to 0.74. Since for some subjects, ear-ECG

TABLE III
COMPARISON OF EAR-ECG SITES, WITH THE VARIANCE IN BRACKETS

	<i>precision</i>	<i>recall</i>	F_1
Mastoid	0.96 (0.02)	0.97 (0.04)	0.97 (0.03)
Concha	0.78 (0.18)	0.76 (0.19)	0.77 (0.18)
In-ear	0.68 (0.16)	0.67 (0.18)	0.69 (0.15)

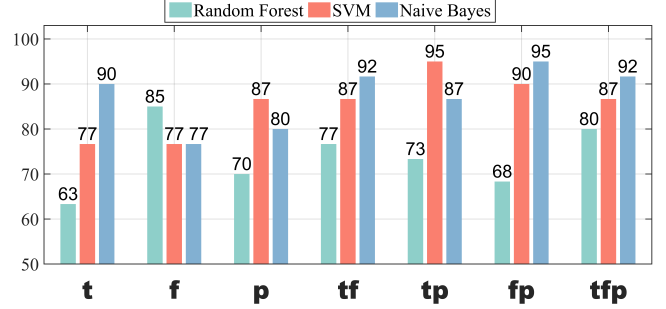


Fig. 6. Classification results based on different feature combinations. On the horizontal axis, **t**, **f**, **p** represent respective classifiers trained with the time domain ECG, frequency domain ECG, and PPG features, defined in Table I. Here, **tf** designates classifiers trained using the combination of time and frequency domain ECG features, **tp** represents classifiers trained using the combination of time domain ECG features and PPG features, and so on.

R-peaks could be extracted from all sites with $F_1 > 0.89$, it was suspected that the observed inconsistent site quality was not physiological. Rather, it could be that the optimum set-up for ear-ECG was harder to achieve compared to the mastoid. Nevertheless, due to the accuracy and consistency of mastoid ECG, the ear-ECG site was used for ECG feature extraction for the rest of this paper.

C. Physiological State Classification

Classification: This work adopted a physiological state classification task to verify the proposed configuration. As Fig. 2 shows, three 5-minute recordings of state (a) normal breathing, (b) slow breathing, and (c) mental exercises (Fig. 2) were available for each subject. Ear-ECG (mastoid) and ear-PPG features shown in Table I were extracted. Different combinations of the extracted features were fed to three classifiers: random forest, SVM, and naive Bayes. The random forest classifier was set to have 250 trees. The SVM used a second-order polynomial kernel. All accuracies were calculated by 49-fold leave-one-out cross-validation. In each fold, data from 6 subjects were used to train the classifiers while data from 1 subject was used for testing. During the 49-fold cross-validation, each subject was left out 7 times.

The naive Bayes achieved the best overall performance (95% using **fp**). It was also observed that adding **p** generally improved classification accuracy. For SVM, **tp** outperformed **t** by 18%, **fp** outperformed **f** by 13%, while **tfp** was similar to **tf**. For naive Bayes, though **tp** had a 3% drop compared to **t**, **fp** outperformed **f** by 18%, **tfp** was similar to **tf**.

The classification of only (a) and (c) without (b) was also verified, in which SVM achieved nearly perfect classification using **tf**, **tp**, or **tfp** features. As hypothesized, the incor-

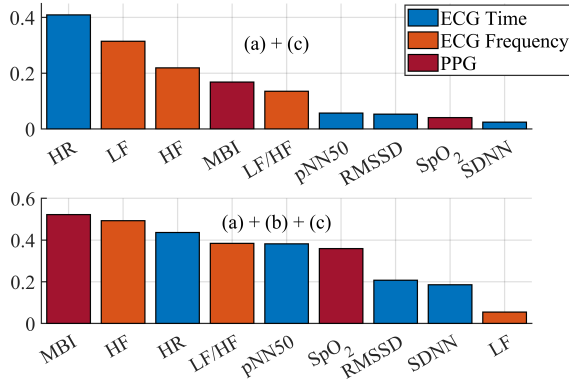


Fig. 7. Permutation feature importance for the random forest classifier. All features (HR, LF, HF, etc.) defined in Table I were used to train the classifier. The top graph was calculated for the binary classification of the sessions: (a) normal breathing and (c) mental exercise. The bottom graph was calculated for the classification of the sessions: (a) normal breathing, (b) slow breathing, and (c) mental exercise.

TABLE IV
CROSS-SUBJECT PERFORMANCE OF THE THREE CLASSIFIERS

	#1	#2	#3	#4	#5	#6	#7
Random Forest	35.3	33.3	100	98.0	33.3	93.7	95.3
SVM	33.3	95.3	100	33.3	95.0	93.7	100
Naive Bayes	66.7	95.3	100	66.7	95.0	93.7	100
Best	66.7	95.3	100	98.0	95.0	93.7	100

poration of both (b) and (c) made the physiological state classification task harder.

Feature importance: To examine whether adding ear-PPG features helped solve the more complicated classification problem of (a), (b) and (c), we calculated the permutation feature importance, which was defined as the decrease in accuracy after randomly permuting the data of one feature. Fig. 7 shows the feature importance without and with (b) controlled slow breathing. An increase in the importance of breathing-related features (MBI and SpO₂) was observed when (b) was present. Although both (b) and (c) cause variance in heartbeat intervals, (b) is related to less stress, and (c) to more stress. Therefore an increase in feature importance was observed for features that reflect stress (HR and SpO₂) and the sympathovagal balance (HF and LF/HF).

Leave-one-out test: To verify the cross-subject robustness of the proposed methods, a leave-one-out test was conducted (Table IV). For column #*i*, data of the *i*-th subject was used as the test set, with data of other subjects as the train set. All **tfp** features were used to train the classifiers. It was observed that the proposed scheme generalized to subjects well, except for subject #1.

V. CONCLUSIONS

This work has demonstrated the feasibility of extracting meaningful HRV features from ear-ECG. Despite the lower SNR, R-peaks were extracted from the mastoid with high accuracy using mMFHT. In terms of ECG R-peak retrieval accuracy, the mastoid has proved better than the concha, while the concha was better than the in-ear site. This work has also proved that the combination of breathing-related features (SpO₂ and MBI) from ear-PPG with ear-

ECG features improved physiological state classification performance, especially in the more complicated case where both (b) controlled slow breathing and (c) mental exercises were present.

Future work will explore the potential for longer-term (24h) HRV monitoring. Compared with standard chest ECG and finger PPG configurations, ear-ECG and ear-PPG could be integrated with headphones and worn for longer periods.

REFERENCES

- [1] W. von Rosenberg, T. Chanwimalueang, V. Goverdovsky, N. S. Peters, C. Papavassiliou, and D. P. Mandic, "Hearables: Feasibility of recording cardiac rhythms from head and in-ear locations," *Royal Society Open Science*, vol. 4, no. 11, p. 171214, 2017.
- [2] G. Hammour, M. Yarici, W. v. Rosenberg, and D. P. Mandic, "Hearables: Feasibility and validation of in-ear electrocardiogram," in *Proc. of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 5777–5780.
- [3] H. J. Davies, I. Williams, N. S. Peters, and D. P. Mandic, "In-Ear SpO₂: A tool for wearable, unobtrusive monitoring of core blood oxygen saturation," *Sensors*, vol. 20, no. 17, p. 4879, 2020.
- [4] G. M. Hammour and D. P. Mandic, "Hearables: Making sense from motion artefacts in ear-EEG for real-life human activity classification," in *Proc. of the 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2021, pp. 6889–6893.
- [5] E. Occhipinti, H. J. Davies, G. Hammour, and D. P. Mandic, "Hearables: Artefact removal in Ear-EEG for continuous 24/7 monitoring," in *Proc. of the IEEE International Joint Conference on Neural Networks (IJCNN)*, 2022, pp. 1–6.
- [6] S. Guler, A. Golparvar, O. Ozturk, and M. K. Yapici, "Ear electrocardiography with soft graphene textiles for hearable applications," *IEEE Sensors Letters*, vol. 6, no. 9, pp. 1–4, 2022.
- [7] K. Budidha and P. Kyriacou, "In vivo investigation of ear canal pulse oximetry during hypothermia," *Journal of Clinical Monitoring and Computing*, vol. 32, no. 1, pp. 97–107, 2018.
- [8] H. J. Davies, P. Bachtiger, I. Williams, P. L. Molyneaux, N. S. Peters, and D. P. Mandic, "Wearable in-ear PPG: Detailed respiratory variations enable classification of COPD," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 7, pp. 2390–2400, 2022.
- [9] P. R. Steffen, T. Austin, A. DeBarros, and T. Brown, "The impact of resonance frequency breathing on measures of heart rate variability, blood pressure, and mood," *Frontiers in Public Health*, vol. 5, p. 222, 2017.
- [10] H.-G. Kim, E.-J. Cheon, D.-S. Bai, Y. H. Lee, and B.-H. Koo, "Stress and heart rate variability: A meta-analysis and review of the literature," *Psychiatry Investigation*, vol. 15, no. 3, pp. 235–245, 2018.
- [11] T. Chanwimalueang, W. von Rosenberg, and D. P. Mandic, "Enabling R-peak detection in wearable ECG: Combining matched filtering and Hilbert transform," in *Proc. of the IEEE International Conference on Digital Signal Processing (DSP)*, 2015, pp. 134–138.
- [12] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, 1985.
- [13] V. Goverdovsky, W. Von Rosenberg, T. Nakamura, D. Looney, D. J. Sharp, C. Papavassiliou, M. J. Morrell, and D. P. Mandic, "Hearables: Multimodal physiological in-ear sensing," *Scientific Reports*, vol. 7, no. 1, pp. 1–10, 2017.
- [14] J. R. Stroop, "Studies of interference in serial verbal reactions," *Journal of Experimental Psychology*, vol. 18, no. 6, p. 643, 1935.
- [15] J. C. Raven and J. H. Court, *Raven's progressive matrices and vocabulary scales*. Oxford psychologists Press Oxford, 1998, vol. 759.
- [16] W. Von Rosenberg, T. Chanwimalueang, T. Adjei, U. Jaffer, V. Goverdovsky, and D. P. Mandic, "Resolving ambiguities in the LF/HF ratio: LF-HF scatter plots for the categorization of mental and physical stress from HRV," *Frontiers in Physiology*, vol. 8, p. 360, 2017.
- [17] MaximIntegrated, "Recommended Configurations and Operating Profiles for MAX30101/MAX30102 EV Kits," <https://pdfserv.maximintegrated.com/en/an/AN6409.pdf>, 2018.
- [18] P. H. Charlton, M. Villarroel, and F. Salguero, "Waveform analysis to estimate respiratory rate," *Secondary Analysis of Electronic Health Records*, pp. 377–390, 2016.