Australasian Physical & Engineering Sciences in Medicine (2018) 41:59–67
https://doi.org/10.1007/s13246-017-0612-9

SCIENTIFIC PAPER

# A principal component analysis based data fusion method for ECG-derived respiration from single-lead ECG

Yue Gao[1] · Hong Yan[1] · Zhi Xu[1] · Meng Xiao[1] · Jinzhong Song[1]

Received: 31 January 2017 / Accepted: 12 December 2017 / Published online: 19 December 2017


## Abstract

An ECG-derived respiration (EDR) algorithm based on principal component analysis (PCA) is presented and applied to derive the respiratory signals from single-lead ECG. The respiratory-induced variabilities of ECG features, P-peak amplitude, Q-peak amplitude, R-peak amplitude, S-peak amplitude, T-peak amplitude and RR-interval, are fused by PCA to yield a better surrogate respiratory signal than other methods. The method is evaluated on data from the MIT-BIH polysomnographic database and validated against a "gold standard" respiratory obtained from simultaneously recorded respiration data. The performance of fusion algorithm is assessed by comparing the EDR signals to a reference respiratory signal, using the quantitative evaluation indexes that include true positive (TP), false positive (FP), false negative (FN), sensitivity (SE) and positive predictivity (PP). The statistically difference is significant among the PCA data fusion method and the EDR methods based on the RR intervals and the RS amplitudes, showing that PCA data fusion algorithm outperforms the others in the extraction of respiratory signals from single-lead ECGs.

**Keywords** Electrocardiogram · ECG-derived respiration · Principal component analysis · Data fusion

## Introduction

Respiration wave is an important physiological signal. It plays an important role in diagnosis and treatment of many diseases, such as obstructive sleep apnea [1, 2]. The respiratory signal is usually obtained with techniques like impedance pneumography, spirometry, or plethysmography. However, such techniques are unmanageable in certain applications such as stress testing, ambulatory monitoring and sleep studies due to the use of some cumbersome devices that may interfere with natural breath [3]. In contrast to traditional methods, electrocardiogram (ECG)-derived respiration (EDR) derives a respiratory signal from the electrocardiogram which is measured noninvasively via a few electrodes [4]. The ECG signal records easily and has less interference with natural breathing which makes the EDR method more suitable for ambulatory monitoring and home monitoring.

✉ Yue Gao
gao_yue1993@126.com

[1] China Astronauts Research and Training Center, No. 26, Beiqing Road, Haidian District, Beijing, China

Respiration influences ECG parameters in two ways, mechanical interaction and respiratory sinus arrhythmia (RSA). Mechanical interaction is caused by two factors: (1) the motion of the electrodes relative to the heart during the respiratory cycle. (2) The thoracic impedance variations caused by the volume changes in the lungs [3, 5]. Respiratory sinus arrhythmia is an important phenomenon in heart rate variability (HRV). It shows how respiration modulates heart rate. The heart rate increases during inspiration and decreases during expiration. The mechanism is that heart rate changes due to respiratory-induced changes to the autonomic nervous system [6].

Due to these factors, a number of EDR algorithms have been described in the literatures. These algorithms provided surrogate respiratory signals based on the respiratory-induced changes of the ECG. The first EDR method based on morphologic variations was proposed by Wang and Calvert [7] which used the vectorcardiogram (VCG) to monitor the respiratory rate and depth. Then some EDR methods based on the amplitude of ECG features like R peak [8, 9], RS [10], T peak [11] or the area under the QRS complex [12] were proposed. A few studies got surrogate respiratory signals through VCG [13] or filtering of the ECG [14–16]. The methods based on RSA derived EDR signals by applying

singular value decomposition (SVD) [17] or S-transform [18] to RR interval series. The methods mentioned above obtained the EDR signals only based on a single piece of information in ECG and this might be improved through data fusion method. Nemati et al. [19] presented a data fusion method based on Kalman-filter to estimate the respiratory rate which proved that data fusion EDR algorithm was more robust to background noise and more accurate. However, this method required multiple physiological sources like ECG, photoplethysmogram, arterial blood pressure and the peripheral arterial tonometry waveforms, and it only derived respiratory rate rather than respiratory waveform. Langley et al. [6] introduced principal component analysis (PCA) to EDR algorithm. They used PCA to look into morphological beat-to-beat variations and applied PCA algorithm to QRS complex and regarded the eigenvectors as EDR signal. They proved that the PCA algorithm outperformed the EDR algorithm based on RS amplitude. Widjaja et al. [5] improved this method by using Kernel PCA to take the nonlinearities in the data into account. These two ways performed well in obtaining EDR signals but they only utilized a part of the respiratory-induced information in ECG. In this article, we combined the ideas of data fusion and PCA to put forward a new EDR algorithm. It should be pointed out that the way we use PCA is completely different to the approach in the two studies mentioned above.

This article is organized as follows. The following section describes the database used in this study. In "ECG-derived respiration" section, the data preprocessing is introduced firstly. Next, the data fusion EDR algorithm based on PCA is presented in detail. The EDR methods based on RR intervals and RS amplitudes are shown at last for comparing. "Comparisons of EDR with standard respiratory signal" section describes the methods for quantitative assessment of respiratory derivation algorithms. The statistical analysis algorithm is given in "Statistical analysis" section. "Result" section presents the result of the proposed method, followed by the "Discussion" and "Conclusion" sections.

## Methods

### Database

The MIT-BIH Polysomnography Database includes 18 continuous records which come from 16 male subjects who suffer from sleep apnoea. Records slp01a and slp01b are segments of one subject's polysomnogram, records slp02a and slp02b are segments of another subject's polysomnogram. The remaining 14 records belong to different subjects. The records are digitized at 250 Hz with 12 bit resolution and continue between 2 and 7 h [20]. The ECG, electromyogram (EMG), electrooculogram (EOG), arterial blood pressure, respiration and arterial blood oxygen saturation are included in the database. The respiration signals are recorded by inductance plethysmography or a nasal thermistor. The ECG and respiration signals were used in this article.

### ECG-derived respiration

#### Data preprocessing

There are three main types of noise in ECG: electromyographical interference, powerline interference and baseline wander. In this article, the electromyographical interference was removed through discrete wavelet transform. Here, we took coif4 as wavelet basis to decompose the ECG into five levels. The soft-threshold equaling to 20% of the mean of the first detail component was employed to this component to remove the electromyographical interference. Next, the powerline interference was handled using average filtering whose width was 10, and the baseline wander was eliminated by polynomial fitting. Then the ECG features detection algorithm based on wavelet [21] was applied on the denoised ECG to get the positons and amplitudes of ECG feature points like P, Q, R, S, T and RR interval.

Then we put forward a histogram denoising method to avoid random noise and guarantee the stability of the EDR signal. The number of the bins in histogram was 100 in this study and more bins have no contribution to the effect of denoising. For example, we got the histogram of all R-peak amplitudes. Then two thresholds were confirmed according to that the ratio of the number of R-peaks, whose amplitudes were located in the region determined by these two thresholds, to all the R-peaks exceeded a fixed value. The ratio was 98% in this article. Finally, the R-peak amplitudes which exceeded the threshold were specified to the threshold. The P-peak amplitudes, Q-peak amplitudes, S-peak amplitudes, T-peak amplitudes and RR intervals were also denoised by this method.

#### PCA data fusion algorithm

(a) Construction of the input matrix

The input matrix $\mathbf{X}$ with dimensions $m \times 6$ consisted of 6 variables including P-peak amplitude (PA), Q-peak amplitude (QA), R-peak amplitude (RA), S-peak amplitude (SA), T-peak amplitude (TA) and RR interval (RR). The 6 variables in each row of $\mathbf{X}$ came from the same cardiac cycle. $m$ was the number of heartbeats included in ECG.

(b) Principal component analysis

The input matrix $\mathbf{X}$ was described as

$$\begin{aligned}\mathbf{X}(\mathrm{n}) &= \left[\boldsymbol{PA}(n)\ \boldsymbol{QA}(n)\ \boldsymbol{RA}(n)\ \boldsymbol{SA}(n)\ \boldsymbol{TA}(n)\ \boldsymbol{RR}(n)\right] \\ &= \left[\boldsymbol{x}_1(n)\ \boldsymbol{x}_2(n)\ \boldsymbol{x}_3(n)\ \boldsymbol{x}_4(n)\ \boldsymbol{x}_5(n)\ \boldsymbol{x}_6(n)\right],\ n = 1, 2, \cdots, m \end{aligned} \quad (1)$$

Firstly, the means of the $\boldsymbol{x}_i$ were removed from $\mathbf{X}$ to get matrix $\boldsymbol{X}'$

$$\mathbf{X}'(\mathrm{n}) = \left[\boldsymbol{x}'_1(n)\, \boldsymbol{x}'_2(n)\, \boldsymbol{x}'_3(n)\, \boldsymbol{x}'_4(n)\, \boldsymbol{x}'_5(n)\, \boldsymbol{x}'_6(n)\right], \quad n = 1, 2, \cdots, m \tag{2}$$

Then the covariance matrix $\mathbf{C}$ of input matrix $\boldsymbol{X}'$ was calculated through

$$\mathbf{C} = \frac{1}{6}\sum_{j=1}^{6} \boldsymbol{x}'_j \boldsymbol{x}'^{T}_j \tag{3}$$

Leading to a 6×6 matrix.

The solution to

$$\lambda \boldsymbol{v} = \mathbf{C}\mathbf{v} \tag{4}$$

yielded the eigenvectors $\mathbf{v} = [\boldsymbol{v}_1\ \boldsymbol{v}_2\ \boldsymbol{v}_3\ \boldsymbol{v}_4\ \boldsymbol{v}_5\ \boldsymbol{v}_6]$and eigenvalues $\boldsymbol{\lambda} = [\lambda_1\ \lambda_2\ \lambda_3\ \lambda_4\ \lambda_5\ \lambda_6]$. The eigenvectors were arranged in order of magnitudes of the corresponding eigenvalues.

Finally, the principal components (PCs) were obtained using

$$\mathbf{PC} = \boldsymbol{X}'\mathbf{V} \tag{5}$$

The first k PCs were retained when the sum of $\lambda_1$ to $\lambda_k$was greater than a ratio of the sum of all the eigenvalues. Here, we fixed the ratio to 60%.

$$\sum_{i=1}^{k} \lambda_i \geq ratio \times \sum_{i=1}^{6} \lambda_i,\ ratio = 60\% \tag{6}$$

The surrogate EDR signal was defined as the sum of the retained PCs.

$$\mathbf{EDR} = \boldsymbol{PC}_1 + \cdots + \boldsymbol{PC}_k \tag{7}$$

The surrogate EDR signal can be also expressed as

$$\mathbf{EDR} = \boldsymbol{X}' \cdot \sum_{i=1}^{k} \boldsymbol{v}_i \tag{8}$$

(c) Interpolation and post-processing

The surrogate EDR signals were interpolated to 250 Hz as for the respiratory signal using cubic spline interpolation for comparison. The baseline of the EDR was obtained by applying a 1-D digital filter with the window of 10 s to the EDR signal. Then the baseline was subtracted from the original EDR signal. At last a Kaiser low-pass filter was used whose cut-off frequency was 0.5 Hz and the passband ripple and stopband attenuation were 0.01 and 0.001 respectively. The reason for choosing 0.5 Hz as cut-off frequency was that normal human's respiration frequency was less than 0.5 Hz [14].

#### EDR methods based on RR interval and RS amplitude

The algorithm proposed in this article was compared with the EDR signals derived from RR intervals and RS amplitudes.

The RR interval time series and RS amplitude series were constructed as

$$\mathbf{RR}(\mathrm{n}) = \boldsymbol{R}(n+1) - \boldsymbol{R}(n)\ n = 1, \cdots, m-1 \tag{9}$$

and

$$\mathbf{RS}(\mathrm{n}) = \boldsymbol{RA}(n) - \boldsymbol{SA}(n)\ n = 1, \cdots, m \tag{10}$$

The $R(n)$ was the time of the n-th R-peaks in the ECG.

Both of the two series were denoised by the histogram method introduced in "Data preprocessing" section. Finally, the surrogate respiratory signals were generated by applying interpolation and post-processing to the denoised series.

## Comparisons of EDR with standard respiratory signal

The assessment method proposed by Mason and Tarassenko [22] was used to compare the EDR signals with the simultaneously recorded reference respiratory signals.

Mason and Tarassenko put forward a quantitative assessment of respiratory derivation algorithms. In their article, the time at which a breath was detected in the EDR signal was compared with the time of the corresponding reference respiratory signal. For breath detection the time window about the reference breath time within which the derived breath should be found was ascertained as 2 s. And every respiration peak or trough was labelled to define a breath beat.

For a detail introduction, a reference respiratory signal and an EDR signal were shown in Fig. 1. There was an antiphase relationship between the two signals. The detailed explanation can be seen in [23]. But not all of the subjects in database were antiphase. In this example, the troughs of the reference signal were served as the breaths and the peaks of the EDR will be considered as breaths. Then the times of the reference respiration breaths will be compared with the times of the EDR breaths [22].

Then the mutually exclusive categories of true positives (TP), false positives (FP), false negatives (FN), sensitivity (SE) and positive predictivity (PP) were employed as the detection statistics. A breath beat in EDR signals that located in the time-match window corresponding to the reference breath will be counted as a TP, and an FP was that had no corresponding breath beat in reference respiration. An FN was a breath beat in the reference respiration that was not matched by a breath in EDR. SE and PP were described as follows:

$$\mathrm{SE} = \frac{TP}{TP + FN} \tag{11}$$

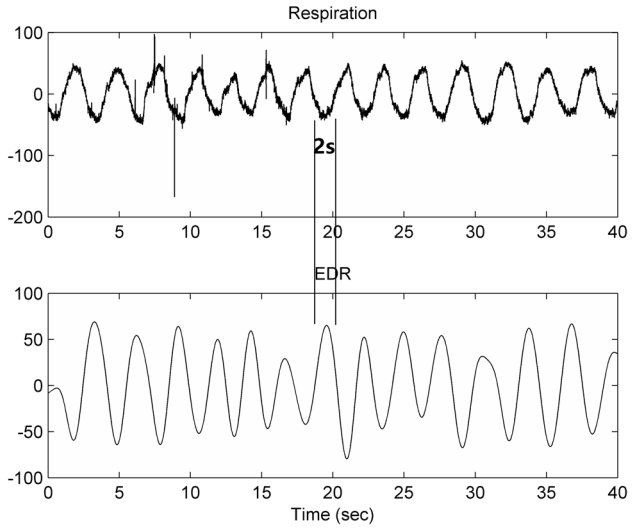



**Fig. 1** A reference respiration and an example EDR. The time interval "2s" marked in figure is the time window about the reference breath time within which the derived breath should be found

$$\mathrm{PP} = \frac{TP}{TP + FP} \tag{12}$$

### Statistical analysis

Friedman Test is used for performance compared in sensitivity and positive predictivity among the PCA data fusion method and those for RR intervals and RS amplitudes. It is a nonparametric statistic that investigates whether multiple samples belong to the same overall. It is considered statistically significant when $p < 0.05$.

## Result

### Visual comparison

The ECG, reference respiratory and EDR signals of subject slp02b are shown in Fig. 2. The relationships among the reference signals and EDR signals are antiphase. All of the EDR methods can recover the surrogate signals, but visual evaluation of EDR-PCA performs better. In the rectangle of the figure, there is a false wave in EDR-RS signal because of the noise in S wave of ECG. However, the performance of the EDR-PCA is still good, which proves that the PCA data fusion method has a better ability to resist noise.

The morphological differences in three EDR signals in Figs. 3 and 4 show that the PCA data fusion method outperforms the other two methods not only in avoiding noise, also the appearance of the EDR-PCA is superior.

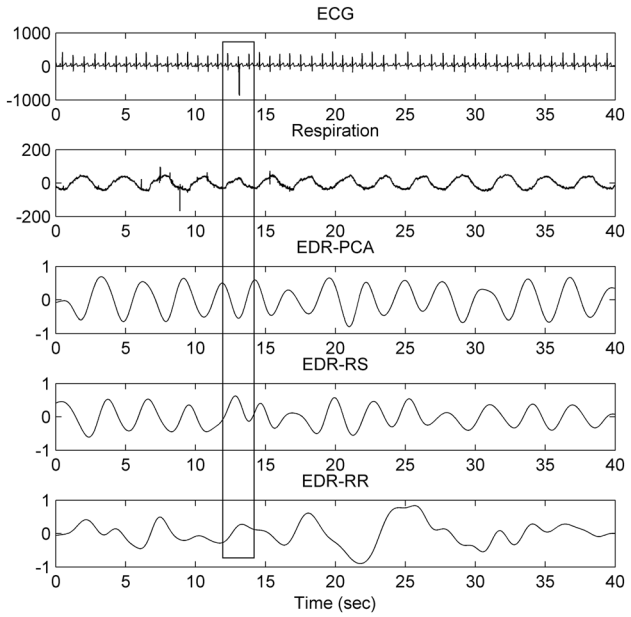



**Fig. 2** ECG and reference respiratory signal from the subject slp02b along with the three EDR signals. The figure shows the data from 0 to 40 s [from top to bottom: the ECG signal; the reference respiratory signal; EDR signal based on PCA data fusion method; EDR signal based on RS amplitudes; EDR signal based on RR intervals]. The rectangular zone shows the ability of three EDR methods to resist noise in ECG. To be convenient for comparison, the EDR signals' amplitudes are normalized

### Quantitative comparison

The comparison results from the three methods for extracting respiration from ECG are shown in Table 1. The number of TPs, FNs, FPs, SE and PP are shown for each record. For quantitative comparison across the whole database, the gross statistics for each algorithm are also calculated.

The comparison of the three different algorithms over all subjects in SE and PP are shown in Figs. 5 and 6. Friedman Test displays significant differences in SE of the three EDR methods ($p < 0.05$), which proves that EDR-PCA is significantly better than EDR-RR and the EDR-RS. However, no statistically significant results are obtained in PP though EDR-PCA is also better than EDR-RS and EDR-RR.

## Discussion

In general, PCA is a data dimension reduction method. In this article, we applied it to EDR analysis and obtained better results than the RS-based and RR-based algorithm. There was a significant difference among the three methods in SE but no significant difference was found in PP. The PCA data fusion algorithm also had the advantage that the EDR-PCA

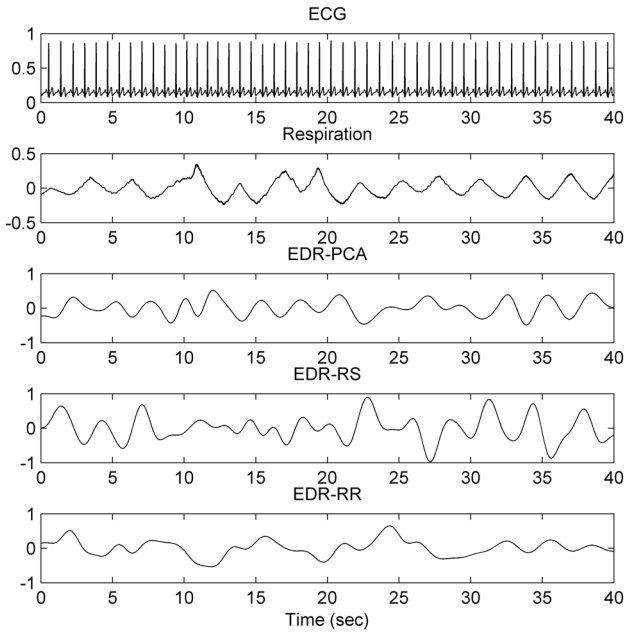



**Fig. 3** ECG and reference respiratory signal from the subject slp66 along with the three EDR signals. The figure shows the data from 0 to 40 s [from top to bottom: the ECG signal; the reference respiratory signal; EDR signal based on PCA data fusion method; EDR signal based on RS amplitudes; EDR signal based on RR intervals]. To be convenient for comparison, the EDR signals' amplitudes are normalized

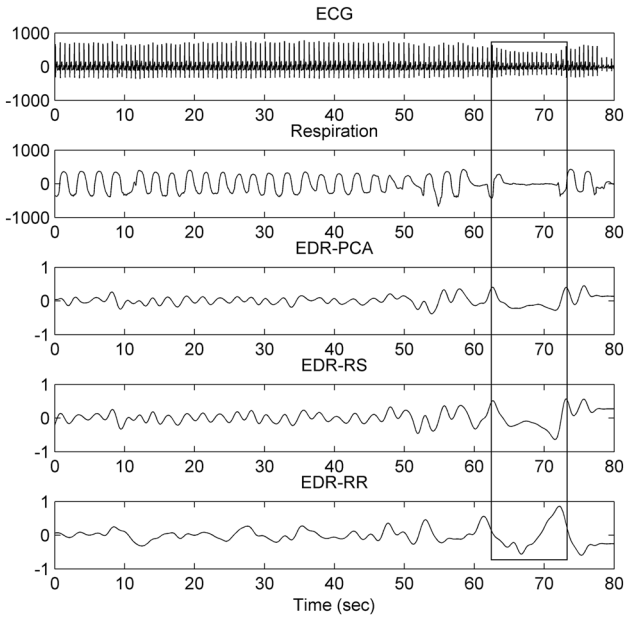



**Fig. 4** ECG and reference respiratory signal from the subject slp16 along with the three EDR signals. The figure shows the data from 0 to 80 s [from top to bottom: the ECG signal; the reference respiratory signal; EDR signal based on PCA data fusion method; EDR signal based on RS amplitudes; EDR signal based on RR intervals]. The rectangular zone shows the ability of three EDR methods to detect the apnoea in subject. To be convenient for comparison, the EDR signals' amplitudes are normalized

was relatively noise free and the appearance of the EDR was more similar to the reference respiration.

In fact, we considered PCA as an approach of data fusion. We can see that, the EDR signal was a linear combination of ECG features such like PA, QA, RA, SA, TA and RR from the formula (2–8). The essence of PCA was to obtain excellent coefficients of the linear combination. So, the method based on RS amplitudes can be considered as a particular case of PCA whose coefficients of RA and SA were 1 and − 1 respectively with the coefficients of other features were 0. The RR intervals algorithm was similar to the method of RS amplitudes. As we know, the theories of EDR methods based on RR intervals and RS amplitudes were RSA and mechanical interaction respectively which were proposed in "Introduction" section. From this point of view, the PCA data fusion method took advantage of all information in single-lead ECG influenced by respiration. Hence, the PCA data fusion method had better robustness and resistance to noise as shown in Figs. 2 and 3. Therefore, it was reasonable that the EDR-PCA outperformed the EDR-RS and EDR-RR.

The coefficients of the linear combination obtained by PCA approach is shown in Table 2. We can see that the coefficients of R, S and T were bigger than the other three coefficients generally. This phenomenon showed that R, S and T sequences were the three most valuable factors to construct EDR signals and proved that these three sequences contained more information of respiration. It was easily to be understood as three articles had obtained the EDR signal based on one of the three sequences separately [8, 10, 11]. The P, Q and RR intervals signals were equally important for modifying the details of EDR signals and resisting noise.

Sleep apnoea occurred in the rectangle region of Fig. 4. It was able to more clearly identify sleep apnoea in EDR-PCA than other EDR signals which showed that the PCA data fusion EDR algorithm had potential in distinguishing sleep apnoea. This will improve the comfort in sleep monitoring for patients without cumbersome devices.

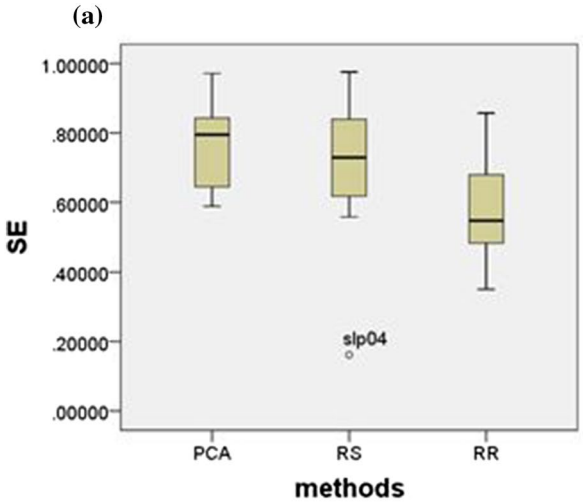
Here, we applied PCA to the whole record. Maybe, we can improve our algorithm by cutting the whole record into several segments to which the PCA will be applied. But this will increase the computation time in turn. The tradeoff between accuracy and time needed more tests. We confirmed the ratio in formula (2–8) to 60% and the ratio in histogram denoising to 98% based on several experiments. The value of the parameters of filters at preprocessing and post-processing sections was fixed in the same way. It required more experiments to get the optimal values for all the records. Nevertheless, the PCA method outperformed the other two methods, even though the ratios were not optimal.

**Table 1** Comparison of all the records

| Record | Method | TP | FP | FN | SE | PP |
|---|---|---|---|---|---|---|
| slp01a | EDR-PCA | 1335 | 142 | 54 | 0.96112 | 0.90386 |
| | EDR-RS | 1355 | 64 | 34 | 0.97552 | 0.9549 |
| | EDR-RR | 944 | 150 | 445 | 0.67963 | 0.86289 |
| slp01b | EDR-PCA | 1829 | 71 | 54 | 0.97132 | 0.96263 |
| | EDR-RS | 1822 | 67 | 61 | 0.9676 | 0.96453 |
| | EDR-RR | 1614 | 136 | 269 | 0.85714 | 0.92229 |
| slp02a | EDR-PCA | 3305 | 634 | 638 | 0.83819 | 0.83905 |
| | EDR-RS | 2736 | 1064 | 1207 | 0.69389 | 0.72 |
| | EDR-RR | 1639 | 1034 | 2304 | 0.41567 | 0.61317 |
| slp02b | EDR-PCA | 2236 | 567 | 463 | 0.82845 | 0.79772 |
| | EDR-RS | 1971 | 771 | 728 | 0.73027 | 0.71882 |
| | EDR-RR | 1189 | 647 | 1526 | 0.43794 | 0.6476 |
| slp03 | EDR-PCA | 5163 | 227 | 962 | 0.84294 | 0.95788 |
| | EDR-RS | 5142 | 243 | 983 | 0.83951 | 0.95487 |
| | EDR-RR | 3626 | 548 | 2499 | 0.592 | 0.86871 |
| slp04 | EDR-PCA | 2516 | 2266 | 1706 | 0.59593 | 0.52614 |
| | EDR-RS | 682 | 3869 | 3540 | 0.16153 | 0.14986 |
| | EDR-RR | 2408 | 1158 | 1823 | 0.56913 | 0.67527 |
| slp14 | EDR-PCA | 3617 | 1298 | 2246 | 0.61681 | 0.73591 |
| | EDR-RS | 3272 | 1769 | 2592 | 0.55798 | 0.64908 |
| | EDR-RR | 3178 | 997 | 2686 | 0.54195 | 0.7612 |
| slp16 | EDR-PCA | 3871 | 1688 | 1586 | 0.70936 | 0.69635 |
| | EDR-RS | 4119 | 1354 | 1338 | 0.75481 | 0.7526 |
| | EDR-RR | 2347 | 1561 | 3110 | 0.43009 | 0.60056 |
| slp32 | EDR-PCA | 2374 | 1286 | 1659 | 0.58864 | 0.64863 |
| | EDR-RS | 2305 | 1306 | 1728 | 0.57153 | 0.63833 |
| | EDR-RR | 1414 | 1719 | 2619 | 0.35061 | 0.45132 |
| slp37 | EDR-PCA | 3668 | 2285 | 2020 | 0.64487 | 0.61616 |
| | EDR-RS | 3515 | 1824 | 2173 | 0.61797 | 0.65836 |
| | EDR-RR | 2745 | 2320 | 2943 | 0.48259 | 0.54195 |
| slp41 | EDR-PCA | 6025 | 867 | 1886 | 0.7616 | 0.8742 |
| | EDR-RS | 5904 | 10,275 | 2007 | 0.7463 | 0.36492 |
| | EDR-RR | 4234 | 975 | 3677 | 0.5352 | 0.81282 |
| slp45 | EDR-PCA | 4230 | 329 | 847 | 0.83317 | 0.92784 |
| | EDR-RS | 3696 | 860 | 1381 | 0.72799 | 0.81124 |
| | EDR-RR | 3585 | 782 | 1492 | 0.70613 | 0.82093 |
| slp48 | EDR-PCA | 4570 | 760 | 359 | 0.92717 | 0.85741 |
| | EDR-RS | 4094 | 1054 | 835 | 0.83059 | 0.79526 |
| | EDR-RR | 2554 | 1250 | 2375 | 0.51816 | 0.6714 |
| slp59 | EDR-PCA | 3207 | 509 | 474 | 0.87123 | 0.86302 |
| | EDR-RS | 3378 | 432 | 303 | 0.91769 | 0.88661 |
| | EDR-RR | 3011 | 411 | 670 | 0.81798 | 0.87989 |
| slp60 | EDR-PCA | 2188 | 1641 | 1400 | 0.60981 | 0.57143 |
| | EDR-RS | 2476 | 1247 | 1112 | 0.69008 | 0.66506 |
| | EDR-RR | 2520 | 1290 | 1068 | 0.70234 | 0.66142 |
| slp61 | EDR-PCA | 3707 | 933 | 762 | 0.82949 | 0.79892 |
| | EDR-RS | 3892 | 726 | 577 | 0.87089 | 0.84279 |
| | EDR-RR | 2878 | 1072 | 1591 | 0.64399 | 0.72861 |
| slp66 | EDR-PCA | 2720 | 976 | 1085 | 0.71485 | 0.73593 |
| | EDR-RS | 2486 | 1041 | 1319 | 0.65335 | 0.70485 |
| | EDR-RR | 2036 | 532 | 1769 | 0.53509 | 0.79283 |

**Table 1** (continued)

| Record | Method | TP | FP | FN | SE | PP |
|---|---|---|---|---|---|---|
| slp67x | EDR-PCA | 663 | 257 | 346 | 0.65709 | 0.72065 |
| | EDR-RS | 585 | 406 | 424 | 0.57978 | 0.59031 |
| | EDR-RR | 558 | 225 | 451 | 0.55302 | 0.71264 |
| Sum | EDR-PCA | 57,224 | 16,736 | 18,547 | | |
| | EDR-RS | 53,430 | 27,066 | 22,342 | | |
| | EDR-RR | 42,480 | 16,807 | 33,317 | | |
| Gross | | | | | 0.75522 | 0.77372 |
| | | | | | 0.70514 | 0.66376 |
| | | | | | 0.56044 | 0.71651 |




(b)

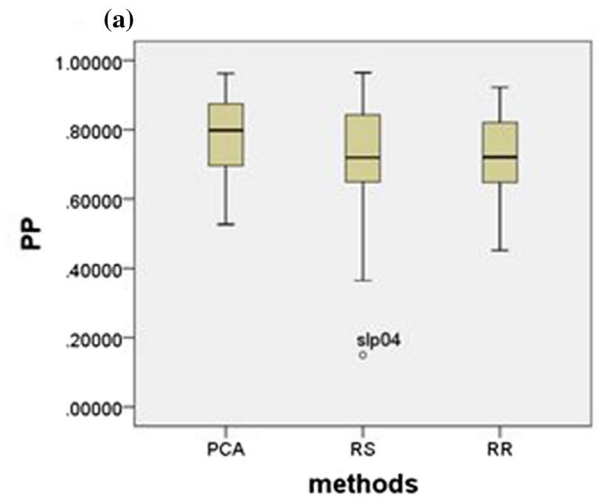
| | Mean Rank |
|---|---|
| PCA | 2.67 |
| RS | 2.17 |
| RR | 1.17 |

(c)

| Test Statistics[a] | |
|---|---|
| N | 18 |
| Chi-Square | 21.000 |
| df | 2 |
| Asymp. Sig. | .000 |

a. Friedman Test

**Fig. 5** Comparison of EDR signals obtained with PCA data fusion method, RS amplitudes and RR intervals over all subjects, expressed in terms of the sensitivity (SE). **a** The boxplots of the three methods. **b** The mean rank of different algorithms in Friedman Test. **c** The test statistics in Friedman Test ($p = 0.000$)




(b)

| | Mean Rank |
|---|---|
| PCA | 2.33 |
| RS | 2.06 |
| RR | 1.61 |

(c)

| Test Statistics[a] | |
|---|---|
| N | 18 |
| Chi-Square | 4.778 |
| df | 2 |
| Asymp. Sig. | .092 |

a. Friedman Test

**Fig. 6** Comparison of EDR signals obtained with PCA data fusion method, RS amplitudes and RR intervals over all subjects, expressed in terms of the positive predictivity (PP). **a** The boxplots of the three methods. **b** The mean rank of different algorithms in Friedman Test. **c** The test statistics in Friedman Test ($p = 0.092$)

## Conclusion

At present, almost all EDR methods obtained the EDR signals based on a single feature from the ECG like R wave or S wave. In this study, we presented a new method for obtaining the ECG-derived respiration signal which combined the ideas of data fusion and PCA technique. The PCA method was applied as a data fusion method to take all of the respiratory-induced variabilities of ECG features into consideration. The results of the algorithm were compared with other two methods. All tests were conducted using MATLAB and the results proved PCA data fusion method to be a promising

**Table 2** The coefficients of the linear combination for all records

| Record | P-coef | Q-coef | R-coef | S-coef | T-coef | RR interval-coef |
|---|---|---|---|---|---|---|
| slp01a | 0.0917 | 0.0009 | 0.9785 | −0.0395 | 0.1806 | −0.0005 |
| slp01b | 0.0934 | 0.0105 | 0.9469 | −0.1789 | 0.2500 | −0.0003 |
| slp02a | 0.1193 | −0.0383 | 0.9853 | −0.0363 | −0.1100 | 0.0003 |
| slp02b | 0.1505 | −0.0300 | 0.9880 | −0.0123 | −0.0075 | 0.0005 |
| slp03 | 0.1023 | −0.0107 | 0.8849 | −0.4458 | −0.0867 | 0.0006 |
| slp04 | 0.2104 | 0.0054 | 0.8118 | −0.1193 | 0.5314 | 0.0007 |
| slp14 | 0.0376 | −0.0280 | 0.9745 | −0.1401 | 0.1690 | 0.0003 |
| slp16 | 0.1197 | 0.0116 | 0.9253 | −0.2462 | 0.2621 | 0.0002 |
| slp32 | 0.0843 | −0.0566 | 0.9621 | −0.2279 | 0.1098 | 0.0000 |
| slp37 | 0.1098 | −0.0322 | 0.9237 | −0.2891 | 0.2238 | 0.0001 |
| slp41 | 0.1885 | 0.1714 | 1.2265 | 0.4208 | 0.5036 | 0.0019 |
| slp45 | 0.0924 | −0.1180 | 0.9794 | 0.1269 | 0.0471 | 0.0000 |
| slp48 | 0.1505 | −0.0808 | 0.9598 | −0.1951 | 0.1077 | 0.0004 |
| slp59 | 0.0407 | 0.0587 | 1.2568 | −0.3647 | 0.5314 | 0.0006 |
| slp60 | −0.0261 | 0.0024 | 0.9828 | −0.1606 | 0.0872 | 0.0000 |
| slp61 | 0.0575 | 0.1924 | 0.7565 | 0.3045 | 0.5429 | 0.0005 |
| slp66 | 0.0182 | 0.0820 | 0.8411 | 0.0843 | 0.5276 | 0.0003 |
| slp67 | −0.0262 | −0.0031 | 0.7132 | 0.1736 | 0.6786 | −0.0006 |

algorithm to obtain surrogate respiratory signal. It was computationally simple, and applicable to single lead ECG. It acquired excellent ECG-derived respiratory signals without the need for additional transducers or hardware which was suitable for home monitoring and ambulatory monitoring.

**Acknowledgements** This work was supported by Advanced Space Medico-Engineering Research Project of China, No. 2014SY54A1102, supported by National Natural Science Foundation of China (61401417), supported by the foundation of China Space Medical Engineering advanced research (2015SY54A0501), supported by State Key Laboratory of Space Medicine Fundamentals and Applications, China Astronaut Research and Training Center (SMFA15A01, SMFA13B03).

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical approval** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

**Informed consent** Informed consent was obtained from all individual participants included in the study.

## References

1. Tiinanen S, Noponen K, Tulppo M, Kiviniemi A, Seppänen T (2015) Ecg-derived respiration methods: adapted ica and pca. Med Eng Phys 37(5):512–517
2. Almazaydeh L, Elleithy K, Faezipour M (2012) Detection of obstructive sleep apnea through ECG signal features. In: IEEE international conference on electro/information technology, vol 3. IEEE, pp 1–6
3. Clifford GD, Azuaje F, McSharry P (2006) Advanced methods and tools for ECG data analysis. Artech House, Inc., Boston
4. Mason L (2002) Signal processing methods for non-invasive respiration monitoring. University of Oxford, Oxford
5. Widjaja D, Varon C, Dorado A, Suykens JA, Van Huffel S (2012) Application of kernel principal component analysis for single-lead-ECG-derived respiration. IEEE Trans Biomed Eng 59(4):1169–1176
6. Langley P, Bowers EJ, Murray A (2010) Principal component analysis as a tool for analyzing beat-to-beat changes in ECG features: application to ECG-derived respiration. IEEE Trans Biomed Eng 57(4):821–829
7. Wang RC, Calvert TW (1974) A model to estimate respiration from vectorcardiogram measurements. Ann Biomed Eng 2(1):47–57
8. Felblinger J, Boesch C (1997) Amplitude demodulation of the electrocardiogram signal (ECG) for respiration monitoring and compensation during MR examinations. Magn Reson Med 38(1):129–136
9. Vargas-Luna JL, Mayr W, Cortés-Ramírez JA (2013) Amplitude modulation approach for real-time algorithms of ecg-derived respiration. Revista Mexicana De Ingenieria Biomedica 35(1):53–69
10. Dobrev D, Daskalov I (1999) Two-electrode telemetric instrument for infant heart rate and apnea monitoring. Med Eng Phys 20(10):729–734
11. Raymond B, Cayton RM, Bates RA, Chappell M (2000) Screening for obstructive sleep apnoea based on the electrocardiogram-the computers in cardiology challenge. In: Computers in cardiology. IEEE, pp 267–270
12. Moody GB, Mark RG, Zoccola A, Mantero S (1985) Derivation of respiratory signals from multi-lead ECGs. Comput Cardiol 12(1985):113–116

13. Leanderson S, Laguna P, Sörnmo L (2003) Estimation of the respiratory frequency using spatial information in the VCG. Med Eng Phys 25(6):501–507
14. Yi WJ, Park KS (2002) Derivation of respiration from ECG measured without subject's awareness using wavelet transform. In: Engineering in medicine and biology. 24th annual conference and the annual fall meeting of the biomedical engineering society EMBS/BMES conference. Proceedings of the second joint, vol 1. IEEE, pp 130–131
15. Labate D, Foresta FL, Occhiuto G, Morabito FC (2013) Empirical mode decomposition vs. wavelet decomposition for the extraction of respiratory signal from single-channel ecg: a comparison. IEEE Sens J 13(7):2666–2674
16. Campolo M, Labate D, Foresta FL, Morabito FC, Lay-Ekuakille A, Vergallo P (2011) ECG-derived respiratory signal using empirical mode decomposition. In: IEEE international workshop on medical measurements and applications proceedings. IEEE, pp 399–403
17. Pilgram B, Di Rienzo M (1993) Estimating respiratory rate from instantaneous frequencies of long team heart rate tracings. In: Computers in cardiology, proceedings. IEEE, pp 859–862
18. Varanini M, De Paolis G, Emdin M, Macerata AAMA, Pola SAPS, Cipriani MACM, Marchesi CAMC (1997) Spectral analysis of cardiovascular time series by the S-transform. In: Computers in cardiology. IEEE, pp 383–386
19. Nemati S, Malhotra A, Clifford GD (2010) Data fusion for improved respiration rate estimation. EURASIP J Adv Signal Process 2010(1), 1
20. Ichimaru Y, Moody GB (1999) Development of the polysomnographic database on CD-ROM. Psychiatry Clin Neurosci 53(2):175–177
21. Yang X, Yan H, Ren Z, Chen J (2010) An ECG feature points detection algorithm based on wavelet. In: National conference on information acquisition and processing (**in Chinese**)
22. Mason CL, Tarassenko L (2001) Quantitative assessment of respiratory derivation algorithms. In: Engineering in medicine and biology society. Proceedings of the 23rd annual international conference of the IEEE, vol 2. IEEE, pp 1998–2001
23. Zhao L, Reisman S, Findley T (1994) Derivation of respiration from electrocardiogram during heart rate variability studies (pp 53–56). IEEE, Long Beach