Journal of Critical Care (2013) **28**, 885.e9–885.e18

ELSEVIER

Journal of Critical Care

Development and validation of a novel fusion algorithm for continuous, accurate, and automated R-wave detection and calculation of signal-derived metrics☆,☆☆,★

Nehemiah T. Liu MS*, Andriy I. Batchinsky MD, Leopoldo C. Cancio MD, William L. Baker Jr BS, Jose Salinas PhD

US Army Institute of Surgical Research, Fort Sam Houston, TX 78234-6315

Keywords:
Signal detection analysis;
Heart rate complexity;
Electrocardiography;
Clinical decision support systems;
Automatic data processing

Abstract

Purpose: Previous studies have shown that heart rate complexity may be a useful indicator of patient status in the critical care environment but will require continuous, accurate, and automated R-wave detection (RWD) in the electrocardiogram (ECG). Although numerous RWD algorithms exist, accurate detection remains a challenge. The purpose of this study was to develop and validate a novel fusion algorithm (Automated Electrocardiogram Selection of Peaks, or AESOP) that combines the strengths of several well-known algorithms to provide a more reliable real-time solution to the RWD problem.

Materials and Methods: This study involved the ECGs of 108 prehospital patient records and 32 ECGs from a conscious sedated porcine model of hemorrhagic shock. The criterion standard for validation was manual verification of R waves.

Results: For 108 human ECG records, the AESOP algorithm overall outperformed each of its component algorithms. In addition, for 32 swine ECG records, AESOP achieved an R-wave sensitivity of 97.9% and a positive predictive value of 97.5%, again outperforming its component algorithms.

Conclusion: By fusing several best algorithms, AESOP uses the strengths of each algorithm to perform more robustly and reliably in real time. The AESOP algorithm will be integrated into a real-time heart rate complexity software program for decision support and triage in critically ill patients.

1. Introduction

Heart rate complexity (HRC) and variability (HRV) tools are used to quantify beat-to-beat changes in the R-to-R interval (RRI), that is, the beat interval between the most prominent peaks (R waves) of the electrocardiogram (ECG) and therefore reflect different physiological factors modulating normal sinus rhythm [1-4]. Recently, a multicenter randomized controlled trial in very-low-birth-weight infants

☆ The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

☆☆ This work was supported by the US Army Institute of Surgical Research (Comprehensive Intensive Care Research Task Area and the Combat Casualty Care Research Program).

★ Disclosures: Drs Batchinsky, Cancio, and Salinas have intellectual property (patent pending, no royalties) dealing with the use of heart rate complexity in combination with other data to identify critical illness in patients. A patent application will be filed by Mr Liu with the US Patent Office for the Automated Electrocardiogram Selection of Peaks algorithm.

* Corresponding author. Tel.: +1 210 539 2459; fax: +1 210 539 5992.
E-mail address: nehemiah.liu@us.army.mil (N.T. Liu).

0883-9441/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jcrc.2013.02.015

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 OCT 2013	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Development and validation of a novel fusion algorithm for continuous, accurate, and automated R-wave detection and calculation of signal-derived metrics.	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)	5d. PROJECT NUMBER
Liu N. T., Batchinsky A. I., Cancio L. C., Baker Jr. W. L., Salinas J.,	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
United States Army Institute of Surgical Research, JBSA Fort Sam Houston, TX	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**10**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

showed that clinical use of HRC-based monitoring in the neonatal intensive care unit may reduce mortality in infants on HRC monitoring vs those without it [1]. Previous studies have shown that HRC and HRV not only are useful for monitoring patients in the critical care environment but may also provide an indication of the need for lifesaving interventions in patients with trauma [2-4]. Because calculation of signal-derived metrics such as HRC and HRV relies on accurate R-wave detection (RWD), a hurdle in using HRC as a new vital sign is that ECG review with manual RWD has been needed, which is not practical for real-time continuous application. Therefore, widespread clinical use of ECG-based monitoring will require continuous, accurate, and automated RWD.

Unfortunately, although the biomedical research community has witnessed a plethora of viable techniques for detecting different fiducial points on the ECG [5,6], RWD remains a practical challenge in many situations [5,6]. Although many published RWD algorithms have performed similarly on ECG recordings from standard databases [5,7-12], from our own experiences and testing (see relevant tables below), they have failed to perform, as well, when confronted with real-time ECG signals in an ambulatory or hospital environment. There may be several reasons for this. First, many RWD algorithms were either designed to process ECGs offline and at specific sampling rates or tested retrospectively without formal real-time verification. Second, standard ECG databases do not always provide the signal morphologies and durations—including setup times, delays, and other temporal factors experienced in the intensive care unit—needed to develop and validate these algorithms. Hence, published RWD algorithms may not have encountered signals with a wide range of morphologies (eg, those described by hemorrhagic shock [HS] or have lost ECG characteristics) and/or different kinds of noise, be it electromechanical, organic, or human related. Lastly, many algorithms have not been adapted to handle data from currently available medical devices and cannot be of clinical use without prior knowledge and tweaking of these algorithms.

The purpose of this study was 2-fold: (1) develop a real-time algorithm that could be integrated into a system for calculating ECG-based metrics and (2) validate it against multiple ECG waveform data sets, including a trauma patient database, an animal protocol consisting of multihour records under conditions of HS and autonomic blockade, and a standard database from Physionet (www.physionet.org). We hypothesized that the fusion of several best, published RWD algorithms could use the strengths of each algorithm and develop a more robust and practical real-time solution to RWD in patients with trauma.

The concept of data fusion, particularly for ECG analysis, has been explored by many authors. For example, since the *Computers in Cardiology Challenge*, which began in 2000, Physionet has helped stimulate the use and development of eclectic fusion methods through competitions [13-22]. However, most of the fusion approaches that were developed for these challenges, be they classifying the quality of ECGs [13] or distinguishing between heart rhythms [15,16,18,20], did not involve multiple RWD algorithms, other than the *PhysioNet/Computers in Cardiology Challenge 2006: QT Interval Measurement* [17]. In this latter case, limited details were provided for developing the fusion algorithm "Meta-6" or describing the fusion process [17]. Although 2 RWD algorithms have been used by Li et al [23,24] for the estimation of heart-related features such as heart rate and blood pressures, their approach involves the fusion of signal quality indices, with scarce discussion on the fusion of RWD outputs and other real-time RWD issues. Likewise, many fusion strategies for RWD have focused on neural networks [25,26] and empirical mode decomposition [27] to adaptively adjust system outputs.

A novelty of our algorithm was use of a "nearest-neighbor" approach to practically fuse the outputs of 4 or more real-time RWD algorithms to calculate a streaming HRC or HRV value. We leveraged not only multiple RWD algorithms but also means, modes, and quality indices from [23,24] to accomplish RWD. Owing to its real-time architecture, our fusion algorithm may be easily integrated into a real-time HRC software program for decision support and triage in trauma patients.

2. Materials and methods

2.1. Algorithm development

Our study commenced with a thorough survey of published RWD algorithms, including wavelet-based techniques [28-30]. We selected a subset of best RWD algorithms for real-time implementation based on their performance against Physionet's Massachusetts Institute of Technology–Beth Israel Hospital (MIT-BIH) Arrhythmia Database and their ease of implementation. This subset consisted of the Pan-Tompkins [7], Hamilton-Tompkins [8], Christov [9], Afonso-Tompkins-Nguyen-Luo [10], and Zong-Moody-Jiang [11] algorithms; the first 2 formed one component in our final detection algorithm.

Because RWD algorithms vary in their mathematical techniques and parameter values, algorithms could portray unique strengths and weaknesses under disparate signal environments [6,23,24]. For this reason, we extended the RWD problem to a problem of fusing multiple detection outputs and classifying them, be they beats or peak-to-peak intervals. For data fusion, we chose an approach similar to a nearest-neighbor algorithm.

Tests against animal ECG waveform records suggested that we could enhance detection performance by adaptively dropping and reselecting individual component algorithms based on their histories of performance. For this, we would need to keep track of which algorithms agreed and disagreed

every time a fusion event took place and also heuristically determine tolerances and time frames for doing so. Nevertheless, our experience proved that the tweaking of automation parameters depended greatly on the ECG sampling rate and signal/noise environment.

Furthermore, ECG degradation caused by noise warranted a need for an index to indicate the ECG signal quality. This index would be an additional output of our final detection algorithm. To assess the signal quality, we used an approach described by Li et al [23,24] and compared the individual performances of component detection algorithms on the waveform. Because different detection algorithms are sensitive to different types of noise, a comparison of how well the components of the final detection algorithm performed within a given time frame provides one estimate of the level of noise in a signal [23,24]. We also addressed the problem of using multiple ECG leads for RWD. Our approach was to run separate instances of the fusion algorithm on each lead's waveform signal and then pick the current instance that yielded the best quality index.

The beat signal quality index (bSQI) for the kth beat was defined by Li et al [23,24] to be the ratio of the number of detected beats agreed by n component algorithms (N_{Matched}) to the number of beats detected by all algorithms (N_{Total}) within a time frame of w seconds:

$$\text{bSQI}(k, w) = N_{\text{Matched}}/N_{\text{Total}}. \quad (1)$$

In other words, whenever a chosen number of component algorithms detected the same R wave of an ECG, a match was recorded. The higher the number of matches within a specified time window, the higher the bSQI. We found this method to be sufficient for comparing the signal qualities of different ECG leads. Because both the MIT-BIH arrhythmia and trauma patient databases involved a single dominant ECG lead (lead II), we used swine ECG waveform data instead to systematically tune the values of N_{Total} and n so that signal quality indices would change only after every N_{Total} beats were detected.

Starting with $N_{\text{Total}} = 10$ and incrementing by multiples of 2 (ie, 12, 14, 16, …) and then by multiples of 5 (ie, 15, 20, 25, …), we determined that $n = 2$, 3 and $N_{\text{Total}} = 30$ yielded indices between 0% and 100% with 3% resolution for sampling frequency of 500 Hz. As such, the procedure to identify pragmatic values of n and N_{Total} was based on heuristics rather than on an optimization function. An index of 0% indicated that the ECG signal might have been noise or not available (disconnected from the subject). For swine with an average heart rate of 120 beats/min, this corresponded to a time frame of roughly 15 seconds.

To make our algorithm compatible with existing in-house developed software, we implemented 4 real-time RWD algorithms and the final data fusion algorithm in Java using the Eclipse Integrated Development Environment and according to key concepts described in Appendix. The 4 algorithms are the Hamilton-Tompkins, Christov, Afonso-Tompkins-Nguyen-Luo, and Zong-Moody-Jiang algorithms (see Fig. 1). In addition, we designed the final data fusion algorithm to be platform independent and operable for both real-time output and ECG waveforms with arbitrary sampling rate. Because all algorithms were implemented in Java for real-time usage, differences in execution times were trivial. Please see Appendix and Refs. [7-11] for more details.

In designing the final data fusion algorithm, which we named the Automated Electrocardiogram Selection of Peaks (AESOP) algorithm, we carefully took into account timing information that meets physiological constraints and performs in real time. Any RRIs less than 240 milliseconds were rejected as outputs. Fusion commences with 4 RWD algorithms as inputs and returns final detected peaks, corresponding times, and a beat signal quality index as outputs in approximate real time. Central to the fusion scheme, the AESOP algorithm selects the end time corresponding to a mode RRI or the RRI closest to a previous averaged decision within a given time frame. In other words, if 2 or more component algorithms detect the same R wave, the AESOP algorithm selects the end time and peak value of the R wave corresponding to the mode RRI's end point, Otherwise, the algorithm selects the end time and peak of the R wave, yielding an RRI closest to the previous averaged 12 RRIs. This implies that one "centroid" (averaged 12 RRIs) is used to compare candidate RRIs and requires no convergence criterion. Optimality is relative to the performance of component algorithms and achieved by selecting the algorithm with nearest-neighbor performance.

Time frames satisfy 2 conditions. First, the time difference between consecutive decisions does not exceed 500 milliseconds, and second, fusion takes place shortly after the detection of 2 candidate beats from the same RWD algorithm. Hence, delays between detected beats are no longer than 500 milliseconds. The AESOP algorithm (and its component algorithms) all required less than 6 seconds to analyze ½-hour record of the MIT-BIH arrhythmia database on an Intel Core Duo Processor E7500 (Intel, Santa Clara, CA) at 2.93 GHz.

To add robustness to the AESOP algorithm, we used a periodic test that checks whether the algorithm should wait for inputs or fuse the current buffer of detected beats. Putting the various fusion strategies together produced a simple yet efficient set of if-then-else rules that incorporates a scheme similar to a nearest neighbor algorithm (see Fig. 2). Again, similarity means that comparisons to a mean RRI are used to select an R wave within a time frame. Because the AESOP algorithm can add additional RWD algorithms for fusion, it is extensible and scalable. We summarize the fusion strategies of the AESOP algorithm in Table 1.

2.2. Animal and clinical validation

We obtained ECG waveform data from multiple ECG databases on Physionet and from protocols approved by our

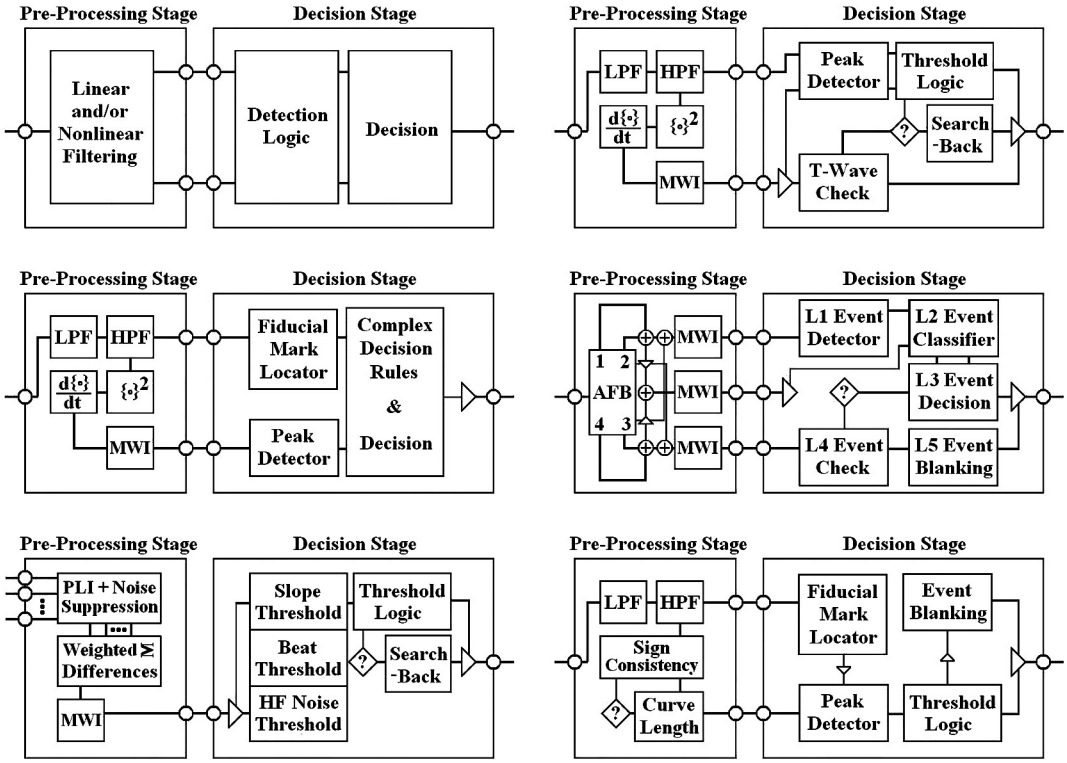

Fig. 1 Diagrams of QRS detection algorithms. A QRS detection algorithm often consists of a preprocessing or feature extraction stage that filters an ECG signal and a decision stage that localizes QRS complexes and other fiducial points. Starting from the top left and moving in a zigzag direction, diagrams depict a standard QRS detection algorithm (top left), the Pan-Tompkins algorithm (top right), the Hamilton-Tompkins algorithm (middle left), the Afonso-Tompkins-Nguyen-Luo algorithm (middle right), the Christov algorithm (bottom left), and the Zong-Moody-Jiang algorithm (bottom right).

institute. The animal study was conducted in compliance with the Animal Welfare Act, the implementing Animal Welfare Regulations, and in accordance with the principles of the Guide for the Care and Use of Laboratory Animals. Where applicable, human studies were approved by the institutional review board at our institute.

During algorithm development, we used ECG lead II records of the MIT-BIH arrhythmia database for identifying

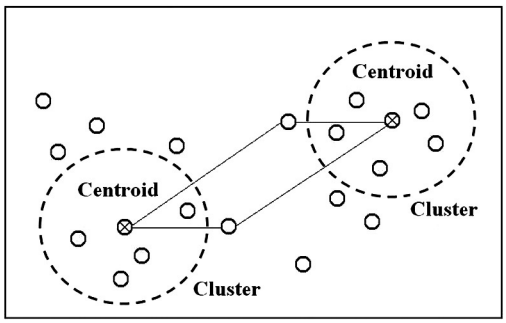

Fig. 2 A nearest-neighbor algorithm in operation. Here, a nearest-neighbor algorithm classifies instances into 1 of 2 groups based on an instance's proximity to a group. To measure proximity, the algorithm first defines a distance measure between an instance and some centroid, which marks the average location of that group. The results of distance comparisons then determine an instance's classification.

Table 1 The automated ECG selection of peaks algorithm

Before fusion
Every quantitative decision must exceed a minimum threshold. Otherwise, no decision occurs within the time frame of fusion.
The time between a previous decision and a candidate decision must exceed a minimum threshold. Otherwise, no decision occurs within the time frame of fusion.
If an input buffer stores a beat from every detector before the detection of 2 consecutive beats from the same detector, then fuse all beats and make a decision within the current time frame.
If an input buffer does not store a beat from every detector before the detection of 2 consecutive or nonconsecutive beats from the same detector, then fuse all beats before the second beat of the same detector and make a decision within the current time frame.
During fusion
If all candidate beats equal, then return the beat as a final decision.
Else if even candidate beats exist and multiple modes exist, then return an arbitrary mode as a final decision.
Else if even or odd candidate beats exist and one mode exists, then return the mode as a final decision.
Else if even or odd candidate beats exist and no mode exists, then return the beat (nearest neighbor) closest to the previous averaged decision (centroid) as a final decision (of a nearest neighbor algorithm)..
Else under a system exception, return no decision.

components for data fusion and testing. We then validated the AESOP algorithm using both trauma patient and animal ECGs. We selected a specific cohort of 108 prehospital patient records from the US Army Institute of Surgical Research Trauma Vitals (TV) database based on ECG availability. Because all data were analyzed post hoc, the study was considered minimal risk, and informed consent was waived. Data records in the TV database include patients with severe trauma (code 2/3) with blunt and penetrating injuries transported from the scene by helicopter service to a level I trauma center in Houston, TX, or San Antonio, TX. Patients were monitored from the scene during transport using a Welch Allyn PIC50 (Welch Allyn, Skaneateles Falls, NY) monitor or a Welch Allyn Propaq 206 (Welch Allyn) monitor. PIC50 data were collected using a computerized personal digital assistant (PDA) attached to the monitor during transport. Data were stored in a nonvolatile memory card in the PDA for use during the study. PIC50 numeric data were stored at a rate of 1 measurement every 3 minutes, coinciding with the patient noninvasive blood pressure measurements. PIC50 waveform data were stored at a rate of 375 Hz. Data from the PDA and flash cards were extracted by research personnel and uploaded to the TV database for analysis. All nonelectronic data were manually recorded on the run sheet from the monitor's screen by emergency medical services medics, then collected on a standardized form and entered into the TV database. Lengths of patient ECGs varied from approximately 15 to 20 minutes.

In addition, we selected 32 swine ECG records from a protocol conducted at our institute. Data records in the animal protocol (A-07-011, Autonomic Blockade protocol) involved healthy, mature, nonsplenectomized, male Sinclair miniature swine with a mean ± SEM weight of 40.4 ± 1.4 kg obtained from the Sinclair Research Center, Inc, Columbia, MO. The protocol studied the effects of atropine and propranolol on the autonomic nervous system during severe HS, which involved removal of 60% of estimated blood volume for 60 minutes. Electrocardiogram waveform data were sampled continuously from the Dräger Evita Infinity V500 ventilator (Dräger, Lübeck, Germany), for the duration of each experiment by using limb ECG leads. Electrocardiograms were acquired at a sampling frequency of 500 Hz and contained electromechanical noise (motion artifacts) as well as organic noise, that is, arrhythmic changes consistent with severe exsanguinations and preterminal and terminal loss of ECG morphology.

Of the 32 records, 10 swine were given vagal blockade (via 0.05 mg/kg intravenous atropine bolus before HS followed by a 0.5-mg kg^{-1} h^{-1} continuous rate infusion during HS) and 12 were given sympathetic blockade (via 0.5 mg/kg propranolol bolus before HS followed by a 0.5-mg kg^{-1} h^{-1} continuous rate infusion during HS). The remaining 10 swine (control) underwent hemorrhage alone. Lengths of records varied from approximately 2 to 5 hours, depending on survival.

The criterion standard for validation was manual verification of R waves, which was accomplished by importing ECG waveform data into WinCPRS software (Absolute Aliens Oy, Turku, Finland), visually analyzing the data, and marking times and positions of all R waves (see Fig. 3). To clarify this process, the WinCPRS software was only used to visualize the ECG waveform data and provide guidance for manually picking times of R waves via mouse clicks on time points of the ECG. After handpicking R waves of all human and animal ECG data records, times and RRIs were written to text files for future reference.

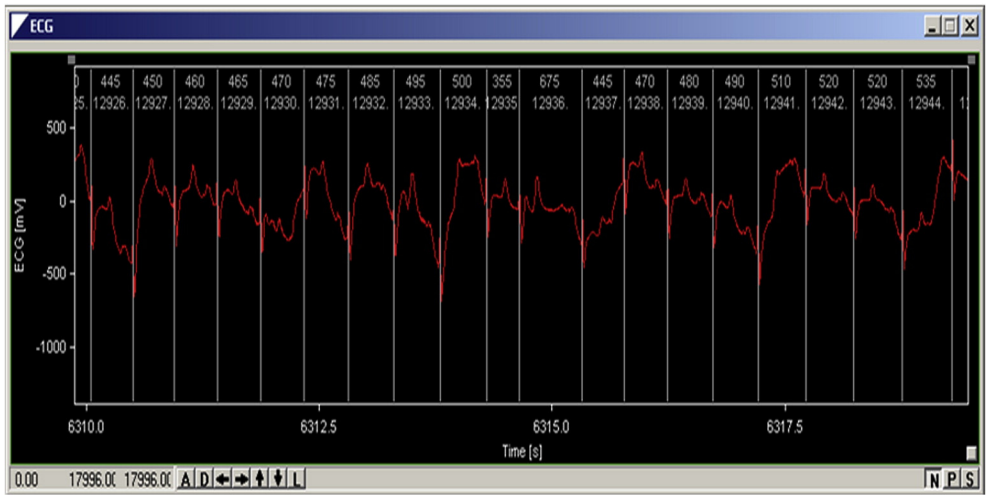

Fig. 3 Example of manually verifying the R waves of an ECG. Manual verification of R waves was accomplished by importing ECG waveform data into WinCPRS software (Absolute Aliens Oy), visually analyzing the data, and manually picking times of R waves via mouse clicks on time points of the ECG. After handpicking R waves of all human and animal ECG data records, times and RRIs were written to text files for future reference.

3. Results

Two benchmark parameters were used to compare the performances of detection algorithms. The sensitivity (Se) and positive predictive value (+P) of the detection algorithm were computed by

$$\mathrm{Se} = \mathrm{TPs}/(\mathrm{TPs} + \mathrm{FNs}), \quad (2)$$

$$+\mathrm{P} = \mathrm{TPs}/(\mathrm{TPs} + \mathrm{FPs}), \quad (3)$$

where TPs is the number of true positives (TPs) or true beats; FNs, the number of false negatives (FNs) or missed beats; and FPs, the number of false positives (FPs) or falsely detected beats. Sensitivity reports the percentage of true beats that were correctly detected by the algorithm, whereas +P reports the percentage of beat detections which were in reality true beats.

For 108 trauma patient records in the US Army's TV database, the AESOP algorithm achieved an R-wave Se of 91.8% and a +P of 92.4%, thereby outperforming each of its component algorithms. In terms of operating points, of 214 823 true beats, AESOP detected 197 257 TPs, 16 603 FPs, and 17 566 FNs. In addition, for 32 swine ECG records from the conscious sedated porcine model of HS, the AESOP algorithm achieved an overall R-wave Se of 97.9% and a +P of 97.5%, thereby outperforming each of its component algorithms in terms of mean Se/+P, that is, tradeoff between Se and +P. Of 815 161 true beats, AESOP detected 797 816 TPs, 20 754 FPs, and 17 345 FNs. Similarly, AESOP performed better than the rest against records from each experimental group of the animal protocol and against all 48 records from the MIT-BIH arrhythmia database (see Tables 2 and 3).

Furthermore, to visually analyze the detection performance of all algorithms, we produced 3 different kinds of plots: (1) plots of RRI sequences over the length of an ECG waveform record, (2) plots of the times of detected R waves superimposed as line segments on an ECG waveform, and (3) histogram plots of the RRI sequences. From these visual analyses, the AESOP algorithm performed desirably. In particular, using the second type of plots mentioned above, the AESOP algorithm merged the 4 individual algorithms in such a way as to remove apparent extremes, most notably, near the beginning of an ECG waveform record (Figs. 4-6).

4. Discussion

From our experiments involving standard test data as well as laboratory data at the Institute of Surgical Research, we observed that the 4 individual RWD algorithms that compose the AESOP algorithm each demonstrated particular strengths and weaknesses, thus reaffirming the use of a fusion algorithm for detecting R waves within an ECG signal, especially in a real-time, ambulatory environment. For specific records, each component algorithm functioned better than the other algorithms according to our implementations.

However, the Hamilton-Tompkins algorithm failed to detect early initial beats and final beats, although it

Table 2 Performance of RWD against porcine model of severe hemorrhage

Algorithm	Group	No.	Verified	TP	FP	FN	Se (%)	+P (%)
AESOP	VB	10	349 755	343 804	7478	5951	98.3	97.9
	SB	12	168 686	165 433	6707	3253	98.1	96.1
	CTL	10	296 720	288 579	6569	8141	97.3	97.8
	Total	32	815 161	797 816	20 754	17 345	97.9	97.5
H-T	VB	10	349 755	343 809	14 294	5946	98.3	96.0
	SB	12	168 686	165 406	10 164	3280	98.1	94.2
	CTL	10	296 720	288 712	8264	8008	97.3	97.2
	Total	32	815 161	797 927	32 722	17 234	97.9	96.1
A-T-N-L	VB	10	349 755	287 000	17 117	62 755	97.3	93.4
	SB	12	168 686	148 562	9687	20 124	97.2	92.8
	CTL	10	296 720	232 156	35 772	64 564	96.0	82.8
	Total	32	815 161	667 718	62 576	147 443	96.7	91.1
C	VB	10	349 755	301 703	10 2031	48 052	86.3	74.7
	SB	12	168 686	131 174	110 433	37 512	77.8	54.3
	CTL	10	296 720	252 000	221 529	44 720	84.9	53.2
	Total	32	815 161	684 877	433 993	130 284	84.0	61.2
Z-M-J	VB	10	349 755	290 987	12 475	58 768	83.2	95.9
	SB	12	168 686	162 531	8857	6155	96.4	94.8
	CTL	10	296 720	273 453	50 644	23 267	92.2	84.4
	Total	32	815 161	726 971	71 976	88 190	89.2	91.0

H-T indicates Hamilton-Tompkins; A-T-N-L: Afonso-Tompkins-Nguyen-Luo; C: Christov; Z-M-J: Zong-Moody-Jiang; VB: vagal blockade; SB: sympathetic blockade; CTL: control.

Table 3 Performance of RWD against ECG databases

Algorithm	Database	TP	FP	FN	Se (%)	+P (%)
AESOP	Animal	797 816	20 754	17 345	97.9	97.5
	Human	197 257	16 603	17 566	91.8	92.2
H-T	Animal	797 927	32 722	17 234	97.9	96.1
	Human	196 896	16 583	17 927	91.7	92.2
A-T-N-L	Animal	287 000	17 117	62 755	97.3	93.4
	Human	178 668	15 863	36 155	83.2	91.9
C	Animal	301 703	102 031	48 052	86.3	74.7
	Human	115 420	28 204	99 403	53.7	80.4
Z-M-J	Animal	726 971	71 976	88 190	89.2	91.0
	Human	143 572	16 661	71 251	66.8	89.6
Manual verification		Total beats				
	Animal	815 161				
	Human	214 823				

H-T indicates Hamilton-Tompkins; A-T-N-L: Afonso-Tompkins-Nguyen-Luo; C: Christov; Z-M-J: Zong-Moody-Jiang.

performed well in many noisy environments. Likewise, the Christov algorithm handled most ECG signals but failed to detect early beats. Moreover, this algorithm blocked occasionally because of major baseline shifts and interruptions within an ECG signal. Like the Hamilton-Tompkins algorithm, the Afonso-Tompkins-Nguyen-Luo algorithm performed acceptably but detected more artifacts than the former. Despite its false detections, though, the Afonso-Tompkins-Nguyen-Luo algorithm detected both early and late beats presumably better than all other algorithms, thereby enhancing fusion. Lastly, the Zong-Moody-Jiang algorithm detected the most artifacts and responded to noise most severely but served as a supplemental check to missed beats.

Through machine learning and data fusion, the AESOP algorithm achieves a balance between the number of FP detections and FN detections, that is, between Se and +P. Importantly, when considering the gamut of signal environments available, AESOP has proven to provide a more reliable solution than its component RWD algorithms. Thus, improving current implementations of AESOP's component algorithms can only improve AESOP's overall performance. Likewise, replacing current components with or adding new reliable components to the AESOP algorithm can increase Se and +P performance.

We recall here that the AESOP algorithm showed an ~8% FP/FN detection rate against prehospital patient ECG waveform data, in which tests were performed by streaming ECG waveform data in real time for RWD. To investigate how this would potentially affect HRC results, we computed the mean HRC values of 2 patient groups (patients who received at least one lifesaving intervention (LSI) vs those who received none) using both detected and manually verified RRI sequences taken from the previous 108 patients with trauma. Although the AESOP-derived HRC values and those derived from manually verified RRI sequences showed a different mean value (paired t-test $P < .01$), the differences between the two patient groups were preserved (see Table 4). Normality was checked for all data via density distributions. For all 2-sample Student t tests between LSI and non-LSI patients, $P < .01$. In a real-time scenario, AESOP's bSQI outputs will help determine the reliability of HRC values, Also, additional filters may be used to adjust HRC values to make this metric meaningful for clinical use.

Interestingly to note, the AESOP algorithm performed better overall against swine ECG records than human ECG records, not because of differences in the sampling rates used to record the ECGs, but mainly because of individual recordings that may have been corrupted by electromechanical noise and artifacts. Likewise, record lengths did not affect AESOP's overall detection performance (although algorithms presented by Pan and Tompkins [7], Hamilton

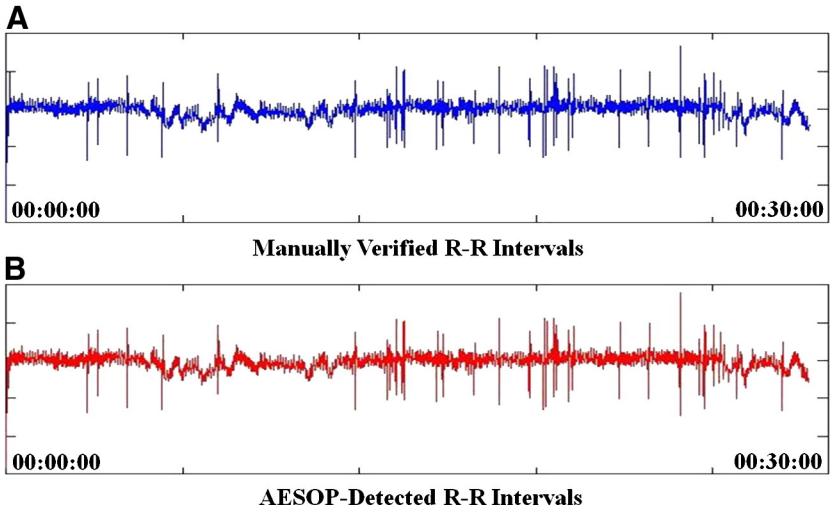

Fig. 4 Example of visually analyzing detection performance. Visual analysis of detection performance included plots of R-R interval sequences over the length of an ECG waveform record, Here, the horizontal axis denotes time in hours and minutes.

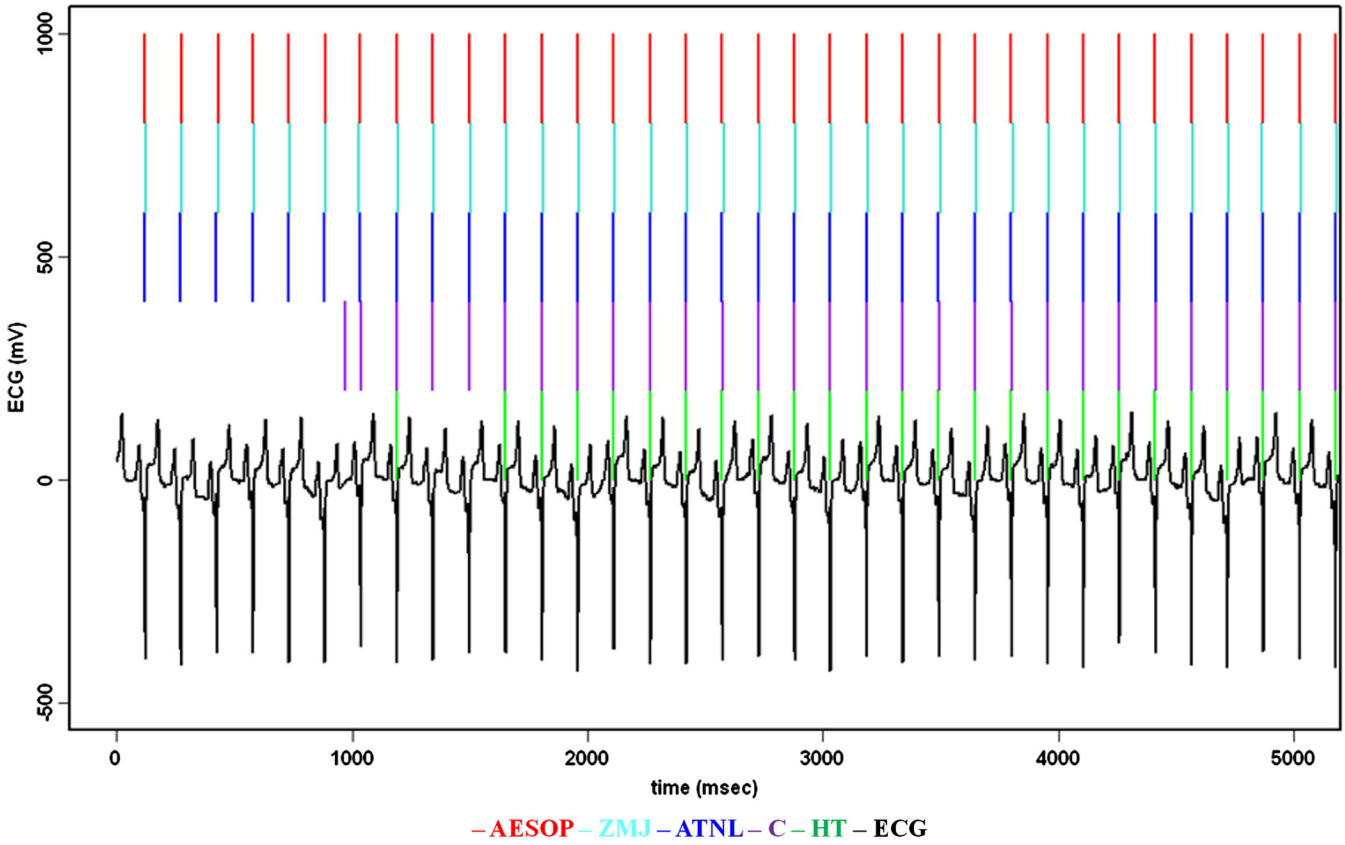

Fig. 5 Example of visually analyzing detection performance. Visual analysis of detection performance included plots of the times of detected R waves superimposed as line segments on an ECG waveform. Here, the horizontal axis denotes time in milliseconds. The fusion algorithm (AESOP) merged the 4 individual algorithms—Zong-Moody-Jiang (ZMJ), Afonso-Tompkins-Nguyen-Luo (ATNL), Christov (C), and Hamilton-Tompkins (HT) algorithms—in such a way as to remove apparent extremes, most notably, near the beginning of the ECG waveform record.

and Tompkins [8], and Suppappola and Sun [12] require some short training time). However, we have observed that human ECGs acquired at a sampling frequency of 182 Hz appear noisier than human ECGs acquired at a sampling frequency of 375 Hz. This difference in quality has resulted in better detection performance when running AESOP against individual records. Although AESOP's component algorithms all use high-pass filters, these filters tend to filter out more high-frequency perturbations of ECGs and less signal components. Thus, high-pass filtering may not affect ECGs sampled at higher frequencies. Although we

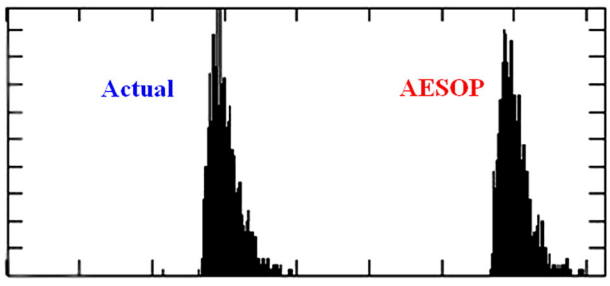

Fig. 6 Example of visually analyzing detection performance. Visual analysis of detection performance included histogram plots of the R-R intervals of an ECG waveform record. Here, the horizontal axis denotes intervals in milliseconds.

Table 4 Comparison of mean HRC values

HRC variable	RWD method	LSI patients (n = 82)	Non-LSI patients (n = 26)
Sample entropy	AESOP	0.8 ± 0.3	1.1 ± 0.4
	H-T	0.9 ± 0.3	1.1 ± 0.4
	A-T-N-L	0.4 ± 0.3	0.5 ± 0.3
	C	0.8 ± 0.3	1.0 ± 0.4
	Z-M-J	0.6 ± 0.4	0.7 ± 0.4
	Manual	0.9 ± 0.3	1.2 ± 0.3
Quadratic sample entropy	AESOP	3.0 ± 0.9	3.5 ± 0.7
	H-T	3.0 ± 1.0	3.5 ± 0.7
	A-T-N-L	4.0 ± 1.1	4.5 ± 0.8
	C	2.9 ± 1.0	3.4 ± 0.8
	Z-M-J	3.7 ± 1.2	4.0 ± 0.5
	Manual	2.7 ± 1.0	3.4 ± 0.8

For paired *t* tests between AESOP and manual verification: $P < .01$. For 2-sample Student *t* tests between LSI and non-LSI patients: $P < .01$. H-T indicates Hamilton-Tompkins; A-T-N-L: Afonso-Tompkins-Nguyen-Luo; C: Christov; Z-M-J: Zong-Moody-Jiang.

conjecture that higher sampling rates will improve the recording quality of ECGs and thus aid RWD, more work will be needed to confirm these assumptions.

Optimizations of AESOP and overall high Se and +P under varying conditions from unperturbed through near-total selective autonomic blockade to HS and organic and electromechanical noise demonstrate its potential application for real-time RWD and HRC monitoring of patients with trauma. The AESOP algorithm will be integrated into a real-time HRC software program for decision support and triage in critically ill and patients with trauma.

Acknowledgments

We thank Corina Necsoiu and Kerfoot Walker III, our colleagues who performed the manual verification of all R waves for 108 trauma patient records in the Trauma Vitals Database and for 32 swine records from a conscious sedated porcine model of severe HS.

Appendix

Although RWD algorithms share a common structure, that is, a preprocessing or feature extraction stage that filters an ECG signal and a decision stage that localizes QRS complexes and other fiducial points, they behave differently because of inherent mathematical techniques and parametric settings (see Fig. 1).

The Pan-Tompkins algorithm serves as a prototype for all threshold-based RWD algorithms. Key concepts used by this algorithm include derivative-based signal processing, integer filters, the adaptation of thresholds using recent signal peaks and noise peaks, a search-back mechanism for finding missed beats, refractory blanking, and T-wave identification [7].

The Hamilton-Tompkins algorithm optimizes as many parameters as possible to construct a complex yet highly efficient set of decision rules that mirrors key inherited concepts [8]. Because of its basis, literature often refers to the Hamilton-Tompkins algorithm as a modified Pan-Tompkins algorithm. Key concepts used by this algorithm include not only those of its parent algorithm but also fiducial mark placement and consistency, mean peak level estimation, baseline shift discrimination, and the optimization of search-back detection thresholds.

Like the Pan-Tompkins algorithm, the Hamilton-Tompkins algorithm preprocesses an ECG signal using a cascaded high-pass filter and low-pass filter, differentiator, nonlinear squaring amplifier, and moving-window integrator (MWI). However, instead of comparing peaks from filtered and integrated signals against dual thresholds, the latter algorithm only compares peaks from the integrated signal against a single threshold. This detection threshold depends on the current mean peak levels according to the following equation:

$$\mathrm{DT} = N_{\mathrm{PL}} + T_{\mathrm{Coeff}} \times (N_{\mathrm{PL}} - \mathrm{QRS}_{\mathrm{PL}}), \tag{4}$$

where DT denotes the detection threshold, T_{Coeff} denotes the threshold coefficient, N_{PL} denotes the mean noise peak level, and $\mathrm{QRS}_{\mathrm{PL}}$ denotes the mean QRS peak level [8].

The Christov algorithm first preprocesses the individual signals with 2 MWIs and after that averages the signals by the following equation:

$$y[n] = \frac{1}{L}\sum_{l=1}^{L} |x_l[n+1] - x_l[n-1]|, \tag{5}$$

where $x_l[n]$ denotes the pre-processed input of lead l at discrete time n, $y[n]$ denotes the averaged signal output, and L denotes the number of leads. In this way, a complex lead is obtained via the above equation by averaging any different (positive and negative) deflection pairs across all primary leads. Then, the algorithm preprocesses $y[n]$ with a third MWI and proceeds like other algorithms to detect the QRS complexes and other characteristic waveforms within an ECG signal. Key concepts used by this algorithm include the combination of 3 independent adaptive thresholds and a search-back mechanism for finding missed beats. In addition, notable strengths of this algorithm include sampling rate and resolution independence and insensitivity to electromyogram and high-frequency noise [9].

The Afonso-Tompkins-Nguyen-Luo algorithm first decomposes an ECG signal into multiple sub-bands for feature extraction. Then, it passes resulting features through a series of decision levels, applying heuristic rules at each level, to accomplish beat detection. Key concepts used by this algorithm include multirate signal processing, signal decomposition into sub-bands using a filter bank, feature extraction, single-channel detection blocks and decision levels, the adaptation of detection strengths using signal and noise histories, and partial refractory blanking. In addition, notable strengths of this algorithm include computational efficiency and resource sharing [10].

Finally, the Zong-Moody-Jiang algorithm transforms a band-limited ECG signal into a set of new values and then compares these values against an adaptive threshold. Key concepts used by this algorithm include the curve length transformation, noise suppression using sign consistency, the adaptation of thresholds, and refractory blanking. In addition, notable strengths of this algorithm include simplicity and the balance between response time and performance [11].

References

[1] Moorman JR, Carlo WA, Kattwinkel J, et al. Mortality reduction by heart rate characteristic monitoring in very low birth weight neonates: a randomized trial. J Pediatr 2011;6:900-6.

[2] Batchinsky AI, Cancio LC, Salinas J, et al. Prehospital loss of R-to-R interval complexity is associated with mortality in trauma patients. J Trauma 2007;63:512-8.
[3] Cancio LC, Batchinsky AI, Salinas J, et al. Heart-rate complexity for prediction of prehospital lifesaving interventions in trauma patients. J Trauma 2008;65:813-9.
[4] Task Force of the European Society of Cardiology, North American Society of Pacing and Electrophysiology: Heart rate variability. Standards of measurement, physiological interpretation and clinical use. Circulation 1996;93:1043-65.
[5] Köhler B-U, Hennig C, Orglmeister R. The principles of software QRS detection. IEEE Eng Med Biol Mag 2002;21:42-57.
[6] Friesen GM, Jannett TC, Jadallah MA, et al. A comparison of the noise sensitivity of nine QRS detection algorithms. IEEE Trans Biomed Eng 1990;37:85-98.
[7] Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Trans Biomed Eng 1985;32:230-6.
[8] Hamilton PS, Tompkins WJ. Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. IEEE Trans Biomed Eng 1986;33:1157-65.
[9] Christov II. Real time electrocardiogram QRS detection using combined adaptive threshold. Biomed Eng Online 2004;3:1-9.
[10] Afonso VX, Tompkins WJ, Nguyen TQ, et al. ECG beat detection using filter banks. IEEE Trans Biomed Eng 1999;46:192-202.
[11] Zong W, Moody GB, Jiang D. A robust open-source algorithm to detect onset and duration of QRS complexes. Comput Cardiol 2003;30:737-40.
[12] Suppappola S, Sun Y. Nonlinear transforms of ECG signals for digital QRS detection: a quantitative analysis. IEEE Trans Biomed Eng 1994;41:397-400.
[13] Silva I, Moody GB, Celi L. Improving the quality of ECGs collected using mobile phones: the PhysioNet/Computing in Cardiology Challenge 2011. Comput Cardiol 2011;38:273-6.
[14] Moody GB. The PhysioNet/Computing in Cardiology Challenge 2010: mind the gap. Comput Cardiol 2010;37:305-8.
[15] Moody GB, Lehman LH. Predicting acute hypotensive episodes: the 10th Annual PhysioNet/Computers in Cardiology Challenge. Comput Cardiol 2009;36:541-4.
[16] Moody GB. The PhysioNet/Computers in Cardiology Challenge 2008: T-wave alternans. Comput Cardiol 2008;35:505-8.
[17] Moody GB, Koch H, Steinhoff U. The PhysioNet/Computers in Cardiology Challenge 2006: QT interval measurement. Comput Cardiol 2006;33:313-6.
[18] Moody GB. Spontaneous termination of atrial fibrillation: a challenge from PhysioNet and Computers in Cardiology 2004. Comput Cardiol 2004;31:101-4.
[19] Moody GB, Jager F. Distinguishing ischemic from non-ischemic ST changes: the Physionet/Computers in Cardiology Challenge 2003. Comput Cardiol 2003;30:235-7.
[20] Moody GB. RR interval time series modeling: the PhysioNet/Computers in Cardiology Challenge 2002. Comput Cardiol 2002;29:125-8.
[21] Moody GB, Goldberger AL, McClennen S, et al. Predicting the onset of atrial fibrillation: the Computers in Cardiology Challenge 2001. Comput Cardiol 2001;28:113-6.
[22] Moody GB, Mark RG, Goldberger AL, et al. Stimulating rapid research advances via focused competition: the Computers in Cardiology Challenge 2000. Comput Cardiol 2000;27:207-10.
[23] Li Q, Mark RG, Clifford GD. Artificial arterial blood pressure artifact models and an evaluation of a robust blood pressure and heart rate estimator. Biomed Eng Online 2009;8:1-15.
[24] Li Q, Mark RG, Clifford GD. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. Physiol Meas 2008;29:15-32.
[25] Hu YH, Tompkins WJ, Urrusti JL, et al. Applications of artificial neural networks for ECG signal detection and classification. J Electrocardiol (Suppl) 1993;26:66-73.
[26] Xue Q, Hu YH, Tompkins WH. Neural-network-based adaptive filtering for QRS detection. IEEE Trans Biomed Eng 1992;39:317-29.
[27] Wu Z, Huang NE. Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal 2009;1:1-41.
[28] Köhler B-U, Hennig C, Orglmeister R. QRS detection using zero crossing counts. Prog Biomed Res 2003;8:138-45.
[29] Li C, Zheng C, Tai C. Detection of ECG characteristic points using wavelet transforms. IEEE Trans Biomed Eng 1995;42:21-8.
[30] Mallat S, Hwang WL. Singularity detection and processing with wavelets. IEEE Trans Inform Theory 1992;38:617-43.