IOP Publishing | Institute of Physics and Engineering in Medicine Physiological Measurement

Physiol. Meas. **37** (2016) 1298–1312 doi:10.1088/0967-3334/37/8/1298

Life-threatening false alarm rejection in ICU: using the rule-based and multi-channel information fusion method

Chengyu Liu[1,2], Lina Zhao[2], Hong Tang[3], Qiao Li[4], Shoushui Wei[2] and Jianqing Li[5]

[1] Departments of Biomedical Informatics, Emory University, Atlanta, GA, USA
[2] School of Control Science and Engineering, Shandong University, Jinan, People's Republic of China
[3] Department of Biomedical Engineering, Dalian University of Technology, Dalian, People's Republic of China
[4] Department of Biomedical Engineering, School of Medicine, Shandong University, Jinan, People's Republic of China
[5] School of Instrument Science and Engineering, Southeast University, Nanjing, People's Republic of China

E-mail: bestlcy@sdu.edu.cn

Received 4 March 2016, revised 19 April 2016
Accepted for publication 26 April 2016
Published 25 July 2016

Abstract

False alarm (FA) rates as high as 86% have been reported in intensive care unit monitors. High FA rates decrease quality of care by slowing staff response times while increasing patient burdens and stresses. In this study, we proposed a rule-based and multi-channel information fusion method for accurately classifying the true or false alarms for five life-threatening arrhythmias: asystole (ASY), extreme bradycardia (EBR), extreme tachycardia (ETC), ventricular tachycardia (VTA) and ventricular flutter/fibrillation (VFB). The proposed method consisted of five steps: (1) signal pre-processing, (2) feature detection and validation, (3) true/false alarm determination for each channel, (4) 'real-time' true/false alarm determination and (5) 'retrospective' true/false alarm determination (if needed). Up to four signal channels, that is, two electrocardiogram signals, one arterial blood pressure and/or one photoplethysmogram signal were included in the analysis. Two events were set for the method validation: event 1 for 'real-time' and event 2 for 'retrospective' alarm classification. The results showed that 100% true positive ratio (i.e. sensitivity) on the training set were obtained for ASY, EBR, ETC and VFB types, and 94% for VTA type, accompanied by the corresponding true negative ratio (i.e. specificity) results of 93%, 81%, 78%, 85% and 50% respectively, resulting in the *score* values of 96.50, 90.70, 88.89, 92.31 and

64.90, as well as with a final *score* of 80.57 for event 1 and 79.12 for event 2. For the test set, the proposed method obtained the *score* of 88.73 for ASY, 77.78 for EBR, 89.92 for ETC, 67.74 for VFB and 61.04 for VTA types, with the final *score* of 71.68 for event 1 and 75.91 for event 2.

Keywords: false alarms rejection, life-threatening arrhythmias, intensive care unit monitor, multi-channel information fusion, rule-based alarm determining, ECG

(Some figures may appear in colour only in the online journal)

1. Introduction

False alarms (FA) in the intensive care unit (ICU) can cause ceaseless and wearisome noise, thus leading to decreased quality of care (Chambrin 2001, Imhoff and Kuhls 2006), such as sleep deprivation, stress for both patients and staff and depressed immune systems. FA rates as high as 86% have been reported (Clifford *et al* 2015), with between only 6% and 40% of ICU alarms having been shown to be true but clinically insignificant (Lawless, 1994). In fact, only 2%–9% of alarms have been found to be important for patient management (Tsien and Fackler 1997). So the intelligent identification method for FAs, especially for life-threatening arrhythmia alarms, plays an important role for clinical application. However, accurately classifying the alarms into true and false ones is still full of challenges.

Various strategies have been employed to deal with the FA problem, including median filtering (Makivirta *et al* 1991), an impulse rejection filter (McNames *et al* 2006, Liu *et al* 2012) and multi-parametric analysis (Zong *et al* 2004, Clifford *et al* 2006, Aboukhalil *et al* 2008). Since the signal quality is a major issue in most cases, the development of signal quality indices (SQIs) can contribute to the improvement of the true/false alarm decision making process. Thus the SQI-based methods for FA rejection have also been widely studied (Li *et al* 2008, Li and Clifford 2012, Behar *et al* 2013).

Previous studies showed that the use of data derived from an independent cardiac-cycle signal might facilitate the FA rejection. The corroboration of alarms using information extracted from a signal highly correlated with the electrocardiogram (ECG) such as a pulsatile waveform that uses an independent sensor to monitor the cardiac cycle, might be able to suppress a large number of false ECG alarms in the ICU. For instance, the validity of an alarm can be determined on the basis of the blood pressure waveform (ABP) when its SQI exceeds a certain threshold; meanwhile the signal quality of the simultaneously recorded ECG is poor. The ABP waveform signal is generated by an independent transducer located away from the torso, exhibits different noise characteristics from an ECG waveform, and is unlikely to contain ECG-related artifacts (except in the case of large body movements of the patient that affect both sensors simultaneously). Therefore, by using information derived from ABP and ECG waveforms, it is likely that true ECG alarms can be effectively corroborated and false ECG alarms suppressed. This has inspired the PhysioNet/CinC Challenge 2015. More detailed description of the background to the Challenge can be found in Clifford *et al* (2015).

Another challenge for FA rejection is that only accurate heart rate estimation has been proven as insufficient to suppress the FAs from ventricular tachycardia (VTA) and ventricular fibrillation arrhythmias in clinical environment (Aboukhalil *et al* 2008, Li *et al* 2014, Salas-Boni *et al* 2014, Fallet *et al* 2015). Thus, in these cases the additional features and methods are required to improve the performance of FA rejection in VTA and ventricular fibrillation arrhythmias.

The aim of the current study was to develop a rule-based and multi-channel information fusion method to reduce the number of FAs and to avoid the suppression of true alarms in the ICU by analysing the simultaneously recorded two channel ECGs, and the possible ABP and/or photoplethysmogram (PPG) signals. Five life-threatening arrhythmias, namely asystole (ASY), extreme bradycardia (EBR), extreme tachycardia (ETC), VTA and ventricular flutter/fibrillation (VFB), were included in the true/false alarm identification. The proposed method consists of five progressively connected steps: signal pre-processing, feature detection and validation, true/false alarm determination for each channel, 'real-time' true/false alarm determination and 'retrospective' true/false alarm determination (if needed). This paper is an extension of our previous work (Liu *et al* 2015) reported in the Physionet/CinC Challenge 2015.

2. Methods

2.1. Dataset

Bedside monitor data leading up to a total of 1250 life-threatening arrhythmia alarm recordings were provided in the Physionet/CinC Challenge 2015 for training and testing the proposed algorithms (Clifford *et al* 2015). The training set contained 750 recordings and the test set contained 500 recordings. Each recording contained two ECG channels and one or two pulsatile waveforms (ABP and/or PPG). All signals were resampled to 12 bit, 250 Hz and had finite impulse response band pass (0.05 to 40 Hz) and main notch filters applied to remove noises.

Five life-threatening arrhythmias (ASY, EBR, ETC, VTA and VFB) were included in the true/false alarms identification. The definitions for the five arrhythmias were summarized as follows (Clifford *et al* 2015):

- ASY: no ECG QRS complex for at least 4 s
- EBR: heart rate lower than 40 beats per min (bpm) for 5 consecutive beats
- ETC: heart rate higher than 140 bpm for 17 consecutive beats
- VTA: 5 or more ventricular beats with heart rate higher than 100 bpm
- VFB: fibrillatory, flutter, or oscillatory waveform for at least 4 s

Each recording lasted for 300 s (short recording) or 330 s (long recording). An alarm was triggered 5 min from the beginning of each record, i.e. the alarm position located at the 300th s of each recording. The alarms in both training and test sets were scored by specialists to verify whether they were true or false alarms. Two events were provided to the Challenge participants to reduce the maximum number of FAs, while avoiding the suppression of true alarms. The recordings in both training and test sets were equally allocated into the two events.

- Event 1—'real-time': using 300 s recordings, i.e. using only the data before the onset of the alarm and no information after that
- Event 2—'retrospective': using 330 s recordings, i.e. using up to 30 s of data after the alarm

2.2. Method description

Figure 1 shows the block diagram of the proposed rule-based and multi-channel information fusion method for identifying the alarms as true or false ones. It consisted of five steps—step 1: signal pre-processing; step 2: feature detection and validation; step 3: true/false alarm determination for each channel; step 4: 'real-time' true/false alarm determination and step 5 (if needed): 'retrospective' true/false alarm determination.

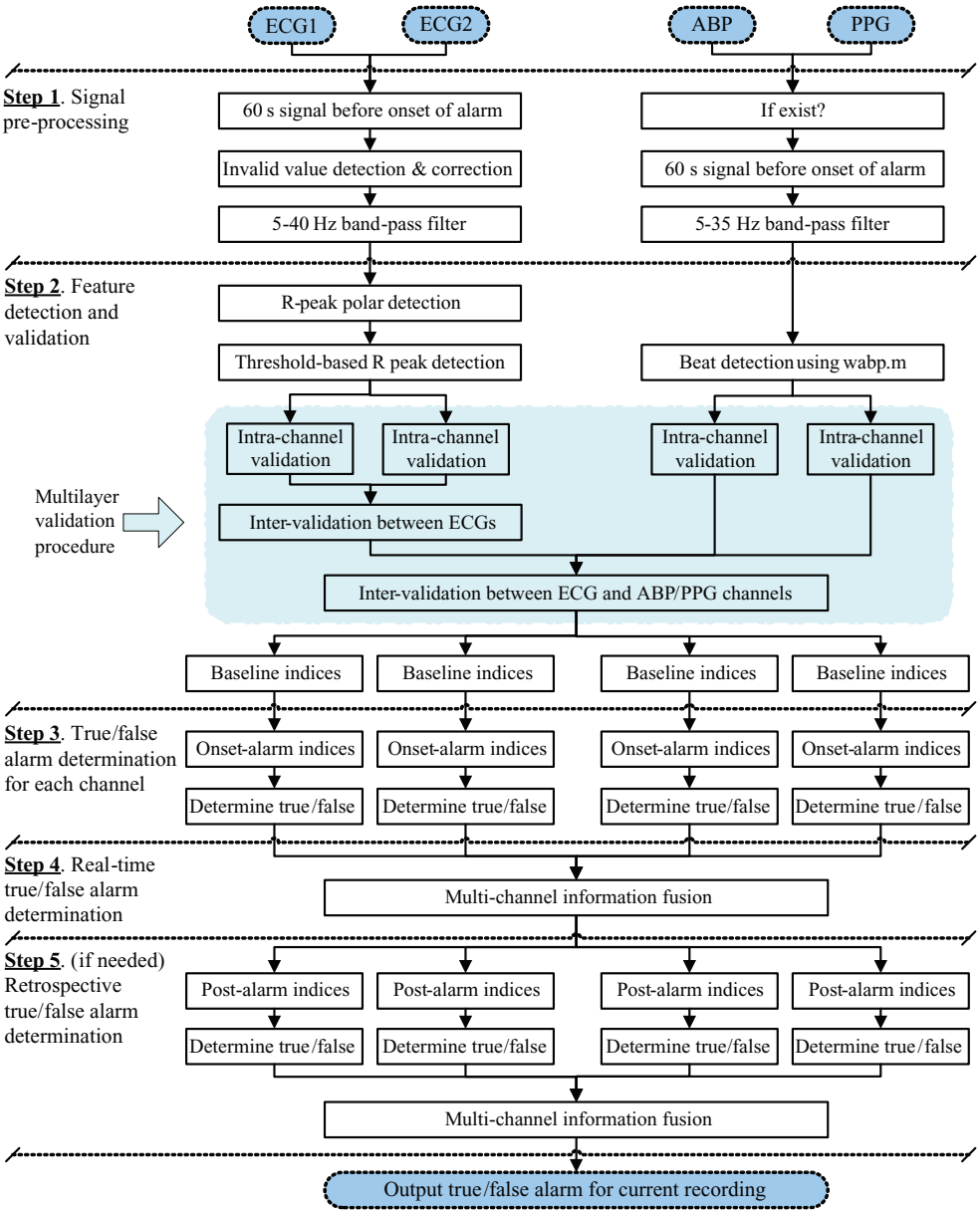

Figure 1. Block diagram of the proposed rule-based and multi-channel information fusion method. Five steps are progressively connected to identify the alarms as true or false ones. Step 5 was only performed on the signal recordings with 330 s length.

2.3. Step 1: signal pre-processing

In step 1, 60 s length signals before the onset of the alarm in each channel were extracted for pre-processing. For the ECG signal, the invalid values of 'Nan' were first corrected using data interpolation. Then a 5–40 Hz band-pass filter was used to further filter the noises. For the ABP/PPG signal, whether the signal existed or not was first checked, then a 5–35 Hz band-pass filter was used for noise rejection.

Table 1. Parameter settings of P_{amp} and P_{time} for the threshold-based R peak detection method.

			Arrhythmia type				
Signal	Parameter	Parameter definition	ASY	EBR	ETC	VTA	VFB
ECG	P_{amp}	Amplitude threshold for determining the candidates of R peaks	0.75	0.75	0.5	0.3	0.5
	P_{time}	Time parameter for determining the candidates of R peaks	0.7	0.7	0.4	0.25	0.5

2.4. Step 2: feature detection and validation

2.4.1. Feature detection. In this section, features (ECG R peaks and pulse feet) were detected for each channel for the filtered 60 s length signal, i.e. looking back 60 s in time from the onset of the alarm.

For the ECG signal, the polar of the R peaks was first detected and the signal was transferred into an inverse signal if the polar was negative. Then a threshold-based detection method was used for the R peaks location and the method was summarized as follows: (1) 60 s ECG signal was first segmented into 10 segments without overlap and the maximum amplitude values in each segment were extracted; (2) the median value of the 10 maximum amplitude values was calculated, and was multiplied by an amplitude parameter P_{amp} as the amplitude threshold; (3) the ECG signal below this amplitude threshold was set as this amplitude threshold and the local maxima in the ECG signal were searched; (4) the local maxima with an interval longer than a time threshold were reserved as the candidates of R peaks, and the time threshold equals the mean of the intervals of the local maxima multiplied by a time parameter P_{time}; (5) finally, the candidates of the R peaks were corrected for the false positive and negative detections, and the final locations of R peaks were re-searched for the local maxima near the candidates. The performances of different combinations of the two parameters P_{amp} and P_{time} were tested using the 750 training recordings for each arrhythmia type and the parameter values with the best results were selected as the final parameter settings shown in table 1. So the settings of amplitude parameter P_{amp} and time parameter P_{time} were different for dealing with different arrhythmia alarm types.

For the ABP/PPG signal, the pulse feet were detected using the open source beat detector of the wabp.m function (Zong *et al* 2003).

2.4.2. Validation of detection results and calculation of baseline indices. After feature detection, a multilayer validation procedure was performed to validate the accuracy of the feature detection results. An intra-channel validation for each channel was first performed and then followed by an inter-channel validation.

For the intra-channel validation, the detected R peaks (or pulse feet) were first checked to meet the minimum number (Min_{num}) criteria. The settings of Min_{num} for each arrhythmia alarm type were given in table 2. Then a distance-matrix-based method was used for further validation and it was summarized as follows: M consecutive locations of R peaks (or pulse feet) from all detected R peaks (or pulse feet) with the minimum standard deviation of RR interval (or foot–foot interval) time series were selected. Then an $M \times M$ distance matrix D was initialized with all 0 elements and was updated using the following rule:

$$D_{i,j} = \begin{cases} 1 & \text{if } \dfrac{\text{Amp}(i)}{\text{Amp}(j)} > \text{Thr}_1 \text{ or } \dfrac{\text{Amp}(j)}{\text{Amp}(i)} > \text{Thr}_1 \\ 0 & \text{else} \end{cases} \quad (1)$$

Table 2. Threshold settings of Min_{num}, Thr_1, Thr_2 and C_{HR} for the multilayer validation procedure.

			Arrhythmia type				
Signal	Parameter	Parameter definition	ASY	EBR	ETC	VTA	VFB
ECG/ ABP/ PPG	Min_{num}	Minimum number of detected R peaks	15	10	30	30	15
ECG	Thr_1	Amplitude ratio threshold	2	2	2	2	2.5
	Thr_2	Threshold for intra-channel validation	0.1	0.1	0.1	0.1	0.25
ABP/ PPG	Thr_1	Amplitude ratio threshold	1.5	1.5	4	2	2.5
	Thr_2	Threshold for intra-channel validation	0.4	0.4	0.2	0.1	0.25
ECG	C_{HR} (bpm)	Pre-set central value of heart rate	75	75	100	100	75

where Amp(i) means the amplitude of the ith R peak or pulse peak (searching in a fixed window after the ith pulse foot), Thr_1 is the amplitude ratio threshold. $D_{i,j} = 1$ means the mismatching of the signal amplitude between the ith and jth R peaks or pulse peaks, indicating the possible detection errors. The current channel will pass intra-channel validation for feature detection only if $\frac{\sum_{i=1}^{M}\sum_{j=1}^{M} D_{i,j}}{M^2}$ is lower than a fixed threshold Thr_2. The settings of thresholds Thr_1 and Thr_2 were also different for different arrhythmia alarm types (see table 2).

Accurately locating the M consecutive R peaks or pulse feet will be verified if the ECG/ABP/PPG channel passed the intra-channel validation procedure. Then the reliable baseline indices within the 60 s time length window before the start of the alarm could be obtained with the reference of these M consecutive R peaks or pulse feet. These baseline indices included:

- *HR_base*: baseline heat rate calculated from M consecutive R peaks or pulse feet, unit: bpm
- *Amp_base*: baseline signal amplitude, i.e. the average value of the M consecutive R peaks or pulse feet
- *Range_base*: baseline signal amplitude range, i.e. the average range of the M consecutive heart cycle signals
- *Template_base*: baseline signal template, i.e. the average of the M consecutive heart cycle signals

In the next step, the inter-channel validation was first performed on two ECG channels. If both ECG channels passed the intra-channel validation, the obtained *HR_base* values were compared to exclude the potential validation errors. The parameter of the central HR (C_{HR}) was pre-set for each arrhythmia alarm type (table 2). If the ratio of two *HR_bases* exceeded the range of [2/3 3/2], the ECG channel with the *HR_base* far from the C_{HR} value failed the validation and this channel was excluded in the following analysis. If only one ECG channel passed the intra-channel validation, this channel was used for the following inter-channel validation. If both ECG channels failed the intra-channel validation, no inter-channel validation procedure between the ECG and ABP/PPG channels was performed. After the inter-channel validation between the ECG channels, the inter-channel validation procedure between the ECG and ABP/PPG channels was performed. The *HR_base* from the ABP/PPG channel was

also compared with that from the validated ECG channel. If the ratio of the *HR_base* values exceeded the range of [2/3 3/2], the ABP/PPG channel was also excluded in the following analysis.

The multiple parameters in table 2 were also trained and optimized using the 750 training recordings for each arrhythmia type. More specifically, the parameters of $\mathrm{Min}_{\mathrm{num}}$ and C_{HR} were determined by a gross observation for both true and false alarm types recording-by-recording. The parameters of Thr_1 and Thr_2 were determined by the grid-search method on a limited range of [1 5] for Thr_1 with a step of 0.5 and [0.05 0.5] for Thr_2 with a step of 0.05.

2.5. Step 3: true/false alarm determination for each channel

In this step, we determined the alarm as true or false for each single ECG/ABP/PPG channel that passed the multilayer validation procedure. The extracted baseline indices were also useful for the determination. The Association for the Advancement of Medical Instrumentation guidelines suggested that each alarm should be raised no later than 10 s from the event onset (2002). So in this study, we used a time window ($T_{\_\mathrm{alarm}}$) of 8 s before the onset of the alarm to analyze ASY, EBR, ETC and VFB arrhythmia types. For the VTA type, we added a 6 s length to $T_{\_\mathrm{alarm}}$, i.e. a 14 s time window was used. This is because some VTA beats in the training set appear before the 8 s $T_{\_\mathrm{alarm}}$ window.

For ASY, EBR and ETC types, the detected features (*R* peaks and pulse feet) in the selected time window ($T_{\_\mathrm{alarm}} = 8$ s) were first verified by comparing the signal amplitude and the baseline indices of *Amp_base* and *Range_base*. The rules were: if the amplitude at the feature point is within the range of [0.5 to 2] times the *Amp_base*, or the signal amplitude range of 0.25 s window centered by the feature point is within the range of [0.5 to 2] times of *Range_base*, the detected feature point was verified as valid. Then the onset-alarm indices were obtained with the reference of the valid feature points:

- *Num_cur*: number of valid feature points
- *HR_cur*: heat rate in the current time window
- *MaxRR_cur*: maximum RR interval in the current time window

The true/false alarm determination rules for each single channel for ASY, EBR and ETC arrhythmia types are shown in figure 2. Figure 3 shows an example for presenting the detailed determination result for recording a582s (ASY, FA).

For the VTA type, ECG channels in the time window $T_{\_\mathrm{alarm}}$ were segmented into four segments with 50% overlap. ABP/PPG channels with VTA arrhythmia and all channels with VFB arrhythmia in the time window $T_{\_\mathrm{alarm}}$ were not segments and were regarded as one segment. For both VTA and VFB types, changes of waveform morphology in each segment were quantified by: (1) calculating the mean correlation degree (*Cor_cur*) between each beat waveform and the constructed signal template *Template_base*; (2) calculating the morphology change rate (*Mor_cur*) by comparing the signal amplitude and the baseline indices of *Amp_base* and *Range_base*. If the amplitude at the feature point was beyond the range of [0.8 to 1.2] times the *Amp_base*, or the signal amplitude range of the 0.25 s window centered by the feature point was beyond the range of [0.8 to 1.2] times the *Range_base*, the current beat was judged as changed morphology.

Mor_cur is the ratio between the changed beat number and the total beat number. The QRS complex width for each beat and maximum of the beat–beat heart rate were also calculated. The added onset-alarm indices were:

- *Cor_cur*: Mean correlation degree in each segment compared with *Template_base*

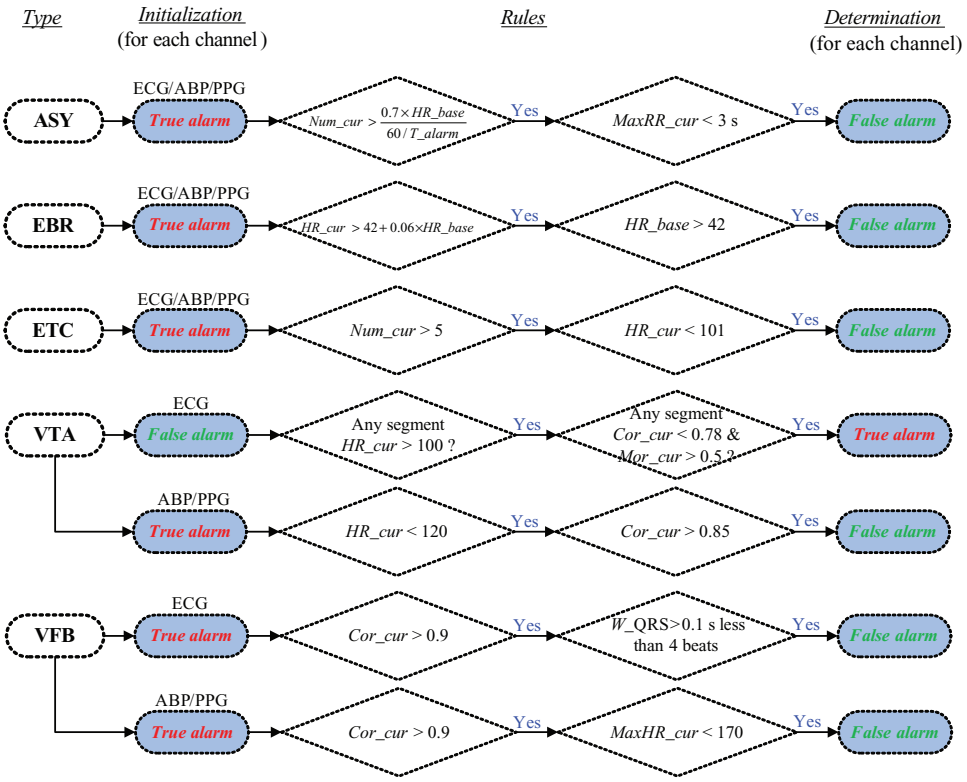

Figure 2. True/false alarm determination rules for each single ECG/ABP/PPG channel signal for the five arrhythmia alarm types.

- *Mor_cur*: Ratio between changed beat number and total beat number in each segment
- *W*_QRS: QRS complex width for each beat
- *MaxHR_cur*: maximum of the beat–beat heart rate

The true/false alarm determination rules for each single channel for VTA and VFB arrhythmia types are also shown in figure 2. Figure 4 shows an example for presenting the detailed determination result for recording v206s (VTA, true alarm).

2.6. Step 4: 'real-time' true/false alarm determination

In this step, the final true/false alarm determination for the whole recording was made using the multi-channel information fusion. The multi-channel information fusion rules were shown in figure 5.

2.7. Step 5 (if needed): 'retrospective' true/false alarm determination

In this step, the first 10 s signals in each ECG/ABP/PPG channel after the alarm were analyzed using the same analysis procedure as shown in step 3 to obtain the post-alarm indices, and thus the true/false alarm determination result for the current channel could be updated based on the results of the post-alarm indices. Then the multi-channel information fusion

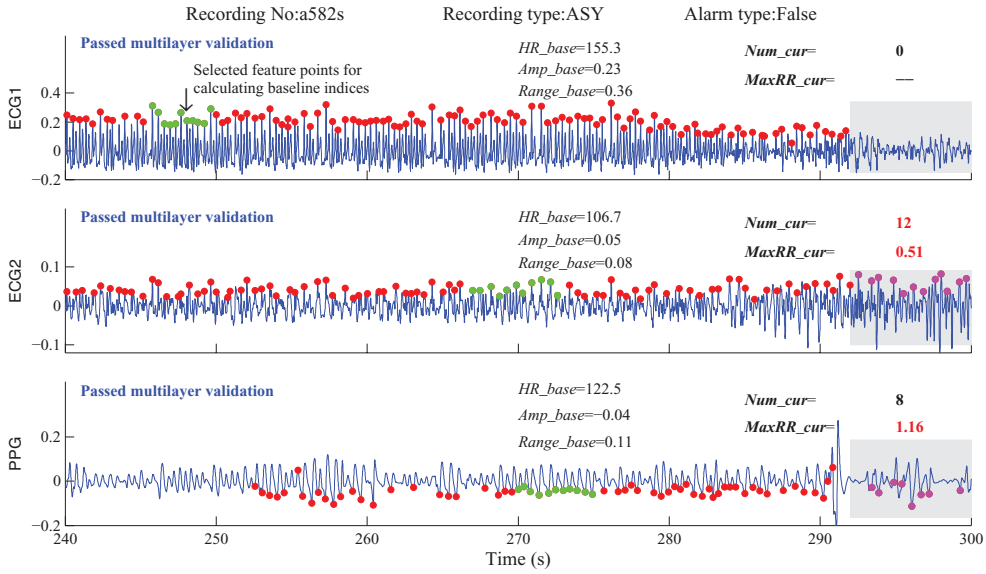

Figure 3. An example for determining a true/false alarm for each channel signal in recording a582s (ASY, false alarm). All three channels have passed the multilayer validation procedure. The detected *R* peaks and pulse feet are marked as a red circle '●', the selected feature points for calculating the baseline indices are marked as a green circle '●' and the valid feature points in the time window $T_{\_alarm}$ (8 s) are marked as a pink circle '●'. The results of the baseline (*HR_base*, *Amp_base* and *Range_base*) and onset-alarm (*Num_cur* and *MaxRR_cur*) indices for each channel are shown at the top of each waveform. The ECG2 channel is determined as FA according to the rules of $Num\_cur > 0.7 \times HR\_base \times T_{\_alarm}/60$ and $MaxRR\_cur < 3$ s (shown in figure 2).

method in step 4 was used again to update the final true/false alarm determination for the whole recording.

2.8. Evaluation indices

The scoring index was a function of the variables: true positives (TP, true alarms classified as true), false positives (FP, false alarms classified as true), false negatives (FN, true alarms classified as false) and true negatives (TN, false alarms classified as false), and was designed to treat FN—genuinely life-threatening events that the program considered unimportant—especially harshly. The corresponding index *score* is defined as (Clifford *et al* 2015):

$$\text{Score} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + 5 \times \text{FN}} \times 100 \tag{2}$$

where the FN detections are penalized five times compared to FP detection.

The indices of the true positive ratio (TPR, i.e. sensitivity) and true negative ratio (TNR, i.e. specificity) were also included in the evaluation, and they are defined as:

$$\text{TPR} = \frac{TP}{TP + FN} \times 100\% \tag{3}$$

$$\text{TNR} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100\% \tag{4}$$

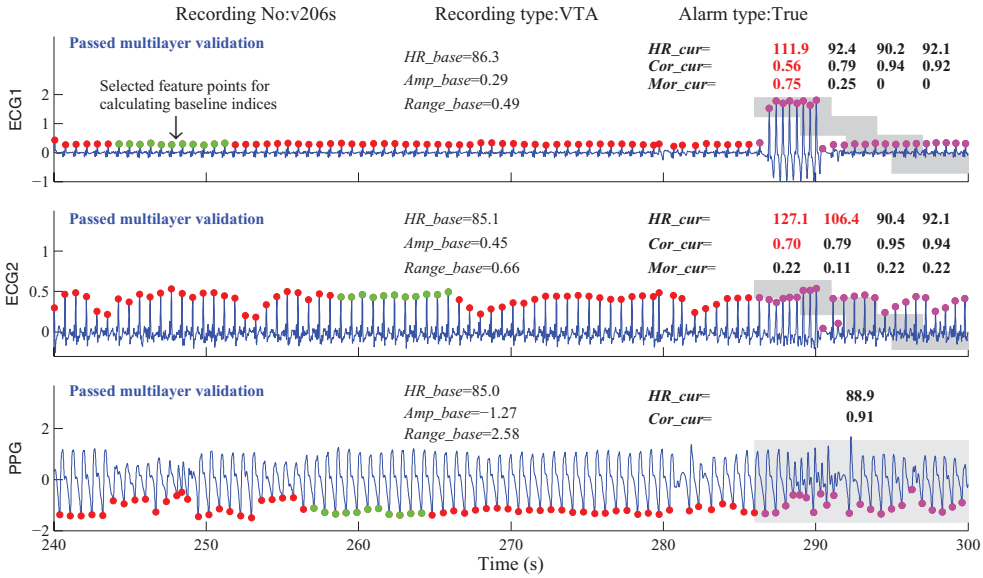

Figure 4. An example for determining a true/false alarm for each channel signal in recording v206s (VTA, true alarm). All three channels have passed the multilayer validation procedure. The detected *R* peaks and pulse feet are marked as a red circle '●', the selected feature points for calculating the baseline indices are marked as a green circle '●' and the valid feature points in the time window $T_{\_alarm}$ (14 s) are marked as a pink circle '●'. Three baseline indices (*HR_base*, *Amp_base* and *Range_base*) for each channel are shown at the top of each panel. ECG signals in the time window $T_{\_alarm}$ are segmented into four segments and the onset-alarm indices (*HR_cur*, *Cor_cur* and *Mor_cur*) for each segment are reported. The results of the onset-indices in ECG1 from the first segment in the time window $T_{\_alarm}$ verify that this channel is a true alarm according to the rules of $HR\_cur > 100$, $Cor\_cur < 0.78$ and $Mor\_cur > 0.5$ (see figure 2).

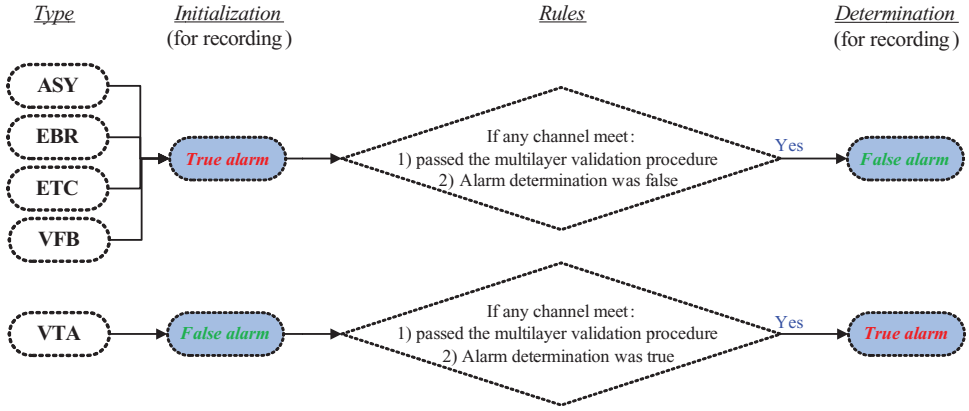

Figure 5. Multi-channel information fusion rules for determining true/false alarms for the whole recording.

Table 3. Evaluation results of the proposed rule-based and multi-channel information fusion method.

	Training set ($N = 750$)							Test set ($N = 500$)		
Arrhythmia type	Number of TP	Number of FN	Number of FP	Number of TN	TPR (%)	TNR (%)	Score	TPR (%)	TNR (%)	Score
ASY	22	0	7	93	**100**	93	96.50	89	93	88.73
EBR	46	0	8	35	**100**	81	90.70	90	91	77.78
ETC	131	0	2	7	**100**	78	88.89	98	60	89.92
VFB	6	0	8	44	**100**	85	92.31	89	69	67.74
VTA	84	5	126	126	94	50	64.90	79	69	61.04
Event 1	151	2	74	148	99	67	**80.57**	89	78	**71.68**
Event 2	138	3	77	157	98	67	**79.12**	93	78	**75.91**

3. Results

3.1. Evaluation results of the proposed method

Table 3 shows the evaluation results of the proposed rule-based and multi-channel information fusion method on both training and test sets. The TPR results from the training set were all 100% for ASY, EBR, ETC and VFB types, and were 94% for the VTA type. The corresponding TNR results were 93%, 81%, 78%, 85% and 50% respectively, resulting in *score* values of 96.50, 90.70, 88.89, 92.31 and 64.90, as well as with a *score* of 80.57 for event 1 and 79.12 for event 2 respectively. The results of the open source entries obtained the *score* values of 88.73 for ASY, 77.78 for EBR, 89.92 for ETC, 67.74 for VFB and 61.04 for VTA types, with the final *score* values of 71.68 for event 1 and 75.91 for event 2.

3.2. Result comparisons among our method, example entries and voting algorithm

Three example entries were provided in the Physionet/CinC Challenge 2015 to serve as a basis for Challenge participants. The results from a voting algorithm were also provided for comparison (Clifford *et al* 2015). This voting algorithm took the top 13 best independent performers' final submissions (judged by the training scores) and voted the submissions together in an un-weighted and trivial manner.

Figure 6 shows the result comparisons of the evaluation indices among our method (on both training and test sets), three example entries (on the test set) and the voting algorithm (on the test set). Our method performed much better than the example entries but had a great gap from the voting algorithm, which reported the highest unofficial s*core* values in the Challenge (*score* 84.26 for event 1 and 87.04 for event 2).

3.3. Evaluation on running time

The algorithm's computational load is important in clinical practice, especially for real-time application. Our method's average and maximum running times are 10.2% and 13.1% of the quota respectively on the training set, and are 10.3% and 12.8% of the quota respectively on the test set, which were within the time limitation of the Challenge.

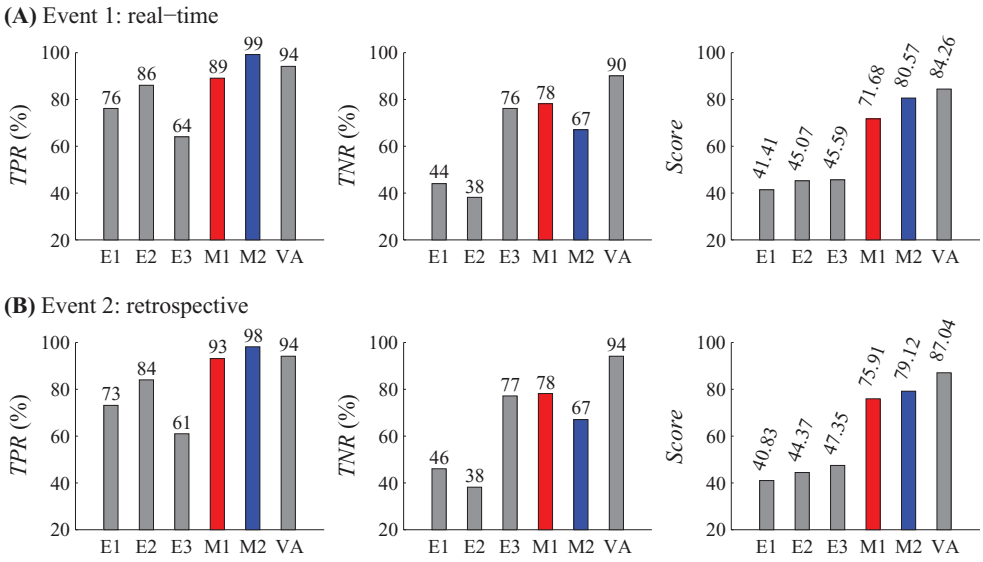

Figure 6. Result comparisons of evaluation indices among our method (on both training and test sets), three example entries (on the test set) and the voting algorithm (on the test set). (A) and (B) are the results for event 1 and event 2 respectively. E1–E3: example entries 1–3 respectively, M1: our results on the test set, M2: our results on the training set, VA: voting algorithm.

4. Discussion

A rule-based and multi-channel information fusion method has been developed to identify the ICU alarms as true or false ones for five life-threatening arrhythmias (ASY, EBR, ETC, VTA and VFB). Its accuracy has been evaluated using a total of 1250 simultaneously collected multi-channel physiological recordings (two ECG channels, and ABP and/or PPG channel) (Clifford *et al* 2015). High sensitivity (i.e. TPR) values of 89% and 93% were achieved for event 1 and event 2 respectively in the official phase of the PhysioNet/CinC Challenge 2015 (Clifford *et al* 2015), accompanied by moderate specificity (i.e. TNR) values of both 78%. For the training set, sensitivity values were even higher and were 99% and 98% for event 1 and event 2 respectively, while specificity values decreased to both 67%.

The unique novelty of the proposed method is that it used a multilayer validation procedure to confirm the accuracy of the detected feature points. Thus reliable baseline indices could be obtained to improve the performance for true/false alarm determination. The function of the validation procedure is similar to the signal quality assessment, which has been proven to play an important, even essential, role in the feature detection of multi-channel physiological signals (Clifford and Moody 2012, Liu *et al* 2014). For the Challenge, both the winners in event 1 (Plesinger *et al* 2015) and event 2 (Fallet *et al* 2015) used the signal quality assessment steps, showing its importance. The multilayer validation procedure ensured the reliable baseline indices, thus improving the performance of the true/false alarm determination that was based on the comparison between baseline and onset-alarm indices, or between baseline and post-alarm indices if retrospective signals were provided. From a review of the works with the top nine scores in the official phase of the Challenge (Ansari *et al* 2015, Antink and Leonhardt 2015, Couto *et al* 2015, Eerikainen *et al* 2015, Fallet *et al* 2015, Kalidas and Tamil 2015, Krasteva *et al* 2015, Liu *et al* 2015, Plesinger *et al* 2015), our method is the only method

including the baseline features from a 60 s time length window before the start of the alarm and performing the comparison between baseline and onset-alarm features, indicating the possibility for developing the personalized determination rules.

Although the scoring function was penalized five times on FN detection compared to FP detection, it is still an important issue that deserves discussion regarding whether the penalization is enough. Zong suggested that rejecting a true alarm was more severe than letting go a false positive alarm five times (Zong 2015). For clinical practice, ICU monitors need 100% sensitivity to all life-threatening arrhythmias. Zong reported that 100% sensitivity was achieved for ASY, EBR, ETC and VFB arrhythmias on the training set. However, the corresponding specificity values of 83%, 83%, 40% and 42% were relatively low. As a comparison, our method also output 100% sensitivity for ASY, EBR, ETC and VFB arrhythmias on the training set, and an output corresponding to the specificity values of 93%, 81%, 78% and 85%. It was also worth noting that the sensitivity values of our method on the test set were not as high as on the training set. One possible reason may be the differences of the signal recordings between the training and test sets.

Moreover, our method has a limited performance on VTA arrhythmia type (sensitivity 79% and specificity 69% for the test set). We noted that the majority of the top performance algorithms (Fallet *et al* 2015, Kalidas and Tamil 2015, Plesinger *et al* 2015) used the spectral characteristics of the ECG signals to detect VTA alarms since the morphology of the QRS complex changed a lot from normal beats to VTA beats, and these morphology changes could impact the huge changes of the spectral characteristics. We used only the amplitude features of the QRS complex but missed the spectral features. Further development by incorporating the spectral analysis methods will be expected to improve the performance of the current method.

With regards to the running time, our method reported the average running times of 10.2% and 10.3% of quota for the training and test sets respectively. Although the time was within the limitation range of the Challenge, it was far longer than the results reported in Zong (2015), which only needed about 0.2% of the quota. So there is still room to improve the running efficiency for the proposed method. On the other hand, it should be noted that step 2 of the proposed method, i.e. the feature detection on the 60 s signals and the multilayer validation procedure for obtaining the baseline indices, occupied most of the running time. This step could be regarded as a pre-learn phase for true/false alarm identification. The constructed baseline features could be reserved as the default values for the segment-by-segment alarm determination in real-time situations. Thus the running time of the code could significantly decrease.

In conclusion, we have proposed a rule-based and multi-channel information fusion method for accurately classifying the true/false alarms for five life-threatening arrhythmias. The proposed method achieved good performances for ASY, EBR and ETC arrhythmias, a moderate performance for VFB arrhythmia and a limited performance for VTA arrhythmia. We have identified the incorporation of the spectral analysis methods to improve the VTA alarm identification, as well as optimizing the algorithm to improve its running efficiency, as our future works.

Acknowledgments

This work was supported by the International Postdoctoral Exchange Programme supported by the National Postdoctoral Management Committee of China, the National Natural Science Foundation of China under Grants of 61201049 and 61571113, and the Excellent Young Scientist Awarded Foundation of Shandong Province in China under Grant BS2013DX029.

References

Aboukhalil A, Nielsen L, Saeed M, Mark R G and Clifford G D 2008 Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform *J. Biomed. Inform.* **41** 442–51

Ansari S, Belle A and Najarian K 2015 Multi-modal integrated approach towards reducing false arrhythmia alarms during continuous patient monitoring: the PhysioNet Challenge 2015 *Computing in Cardiology* ed A Murray (Nice: IEEE) pp 1181–4

Antink C H and Leonhardt S 2015 Reducing false arrhythmia alarms using robust interval estimation and machine learning *Computing in Cardiology* ed A Murray (Nice: IEEE) pp 285–8

Association for the Advancement of Medical Instrumentation 2002 *American National Standard (ANSI/AAMI EC13:2002) Cardiac Monitors, Heart Rate Meters, and Alarms* (Arlington, VA: Association for the Advancement of Medical Instrumentation)

Behar J, Oster J, Li Q and Clifford G D 2013 ECG signal quality during arrhythmia and its application to false alarm reduction *IEEE Trans. Biomed. Eng.* **60** 1660–6

Chambrin M C 2001 Review: alarms in the intensive care unit: how can the number of false alarms be reduced? *Crit. Care* **5** 184–8

Clifford G D, Aboukhalil A, Zong W, Sun J, Moody G B and Mark R G 2006 Using the blood pressure waveform to reduce critical false ECG alarms *Computing in Cardiology* ed A Murray (Valencia: IEEE) pp 829–32

Clifford G D and Moody G B 2012 Signal quality in cardiorespiratory monitoring *Physiol. Meas.* **33** E1–5

Clifford G D, Silva I, Moody B, Li Q, Kella D, Shahin A, Kooistra T, Perry D and Mark R G 2015 The PhysioNet/Computing in Cardiology Challenge 2015: reducing false arrhythmia alarms in the ICU *Computing in Cardiology* ed A Murray (Nice: IEEE) pp 273–6

Couto P, Ramalho R and Rodrigues R 2015 Suppression of false arrhythmia alarms using ECG and pulsatile waveforms *Computing in Cardiology* ed A Murray (Nice: IEEE) pp 749–52

Eerikainen L M, Vanschoren J, Rooijakkers M J, Vullings R and Aarts R M 2015 Decreasing the false alarm rate of arrhythmias in intensive care using a machine learning approach *Computing in Cardiology* ed A Murray (Nice: IEEE) pp 293–6

Fallet S, Yazdani S and Vesin J M 2015 A multimodal approach to reduce false arrhythmia alarms in the intensive care unit *Computing in Cardiology* ed A Murray (Nice: IEEE) pp 277–80

Imhoff M and Kuhls S 2006 Alarm algorithms in critical care monitoring *Anesth. Analg.* **102** 1525–37

Kalidas V and Tamil L S 2015 Enhancing accuracy of arrhythmia classification by combining logical and machine learning techniques *Computing in Cardiology* ed A Murray (Nice: IEEE) pp 733–6

Krasteva V, Jekova I, Leber R, Schmid R and Abacherli R 2015 Validation of arrhythmia detection library on bedside monitor data for triggering alarms in intensive care *Computing in Cardiology* ed A Murray (Nice: IEEE) pp 737–40

Lawless S T 1994 Crying wolf: false alarms in a pediatric intensive care unit *Crit. Care Med.* **22** 981–5

Li Q and Clifford G D 2012 Signal quality and data fusion for false alarm reduction in the intensive care unit *J. Electrocardiol.* **45** 596–603

Li Q, Mark R G and Clifford G D 2008 Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter *Physiol. Meas.* **29** 15–32

Li Q, Rajagopalan C and Clifford G D 2014 Ventricular fibrillation and tachycardia classification using a machine learning approach *IEEE Trans. Biomed. Eng.* **61** 1607–13

Liu C Y, Li L P, Zhao L N, Zheng D C, Li P and Liu C C 2012 A combination method of improved impulse rejection filter and template matching for identification of anomalous intervals in RR sequences *J. Med. Biol. Eng.* **32** 245–50

Liu C Y, Li P, Di Maria C, Zhao L N, Zhang H G and Chen Z Q 2014 A multi-step method with signal quality assessment and fine-tuning procedure to locate maternal and fetal QRS complexes from abdominal ECG recordings *Physiol. Meas.* **35** 1665–83

Liu C Y, Zhao L N and Tang H 2015 Reduction of false alarms in intensive care unit using multi-feature fusion method *Computing in Cardiology* ed A Murray (Nice: IEEE) pp 741–4

Makivirta A, Koski E, Kari A and Sukuvaara T 1991 The median filter as a preprocessor for a patient monitor limit alarm system in intensive care *Comput. Methods Programs Biomed.* **34** 139–44

McNames J, Thong T and Aboy M 2006 Impulse rejection filter for artifact removal in spectral analysis of biomedical signals *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **1** 145–8

Plesinger F, Klimes P, Halamek J and Jurak P 2015 False alarms in intensive care unit monitors: detection of life-threatening arrhythmias using elementary algebra, descriptive statistics and fuzzy logic *Computing in Cardiology* ed A Murray (Nice: IEEE) pp 281–4

Salas-Boni R, Bai Y, Harris P R, Drew B J and Hu X 2014 False ventricular tachycardia alarm suppression in the ICU based on the discrete wavelet transform in the ECG signal *J. Electrocardiol.* **47** 775–80

Tsien C L and Fackler J C 1997 Poor prognosis for existing monitors in the intensive care unit *Crit. Care Med.* **25** 614–9

Zong W 2015 Reduction of false critical ECG alarms using waveform features of arterial blood pressure and/or photoplethysmogram signals *Computing in Cardiology* ed A Murray (Nice: IEEE) pp 289–92

Zong W, Heldt T, Moody G and Mark R 2003 An open-source algorithm to detect onset of arterial blood pressure pulses *Computing in Cardiology* ed A Murray (Thessaloniki Chalkidiki: IEEE) pp 259–62

Zong W, Moody G B and Mark R G 2004 Reduction of false arterial blood pressure alarms using signal quality assessment and relationships between the electrocardiogram and arterial blood pressure *Med. Biol. Eng. Comput.* **42** 698–706