Subject-Optimized Feature Selection for Accurate Classification of Cardiac Beats

Piotr Augustyniak

AGH University of Science and Technology, Krakow, Poland

Abstract

This paper presents a new method of subject-dependent optimization of ECG feature set for heart beats classification. The algorithm learns from randomly selected strips of the signal and preliminarily classifies the example beats. After an approval from a human operator, these classified beats are distributed to a learning and testing sets, and a genetic algorithm with aggressive mutation is used to select features with best discriminative power. Thanks to the integer representation of features in few genes only, the initial feature space of 57 elements is limited by the algorithm to 3 – 5 features optimized for a particular subject.

The proposed method was tested on MIT-BIH Arrhythmia database providing reference beat types for each record. We implemented algorithms calculating 57 different parameters of the beat (mainly focused on the QRS), based on shape, acceleration, area, length etc. and used kNN and SVM as classification methods. We built the learning set out of 15 strips of 10s length and assume that the feature set contains maximum of 5 elements.

Comparing the results to the classification based on reference minimum correlation-based selection of features we observed significant reduction of misclassified beats ratio (for SVM and 3 features from 2.7% to 0.7% in average).

1. Introduction

Beat classification belongs to fundamental procedures in ECG signal processing chains aimed at automatic interpretation. It is a difficult task due to high subject-related variability of the features, moreover, numerous applications require real time processing regime. The aim of heart beat classification is usually twofold: diagnostic, that discloses the origin of the stimulus and represents the activation of conduction pathways and technical, that enables conditional processing of certain beat types (e.g. sinus rhythm for exercise tests or heart rate variability studies). The rhythm classification uses various temporal and shape-related parameters that represent irregularities in generation and conduction of the stimulus and assigns each heart beat to one of morphology categories defined a priori on the physiological background. The reference MIT-BIH Arrhythmia Database [1] distinguishes 41 beat types, although only few of them are detected in a single record. The technically-oriented classification also use various temporal and voltage-derived factors, but yields only four main beat categories usually named: supraventricular, ventricular, other and non-QRS.

The problem of accurate and fast classification of heart beats was a scientific and engineering challenge since early years of computerized electrocardiography. In recent years O'Dwyer et al. [2] analyzed the suitability of different time and amplitude-based features for heart beat classification. Chang et al. [3] compared four most common metrics to measure the similarity of two sections of the ECG signal. An interesting QRS classification method based solely on ECG sample values was proposed by Lemay et al. [4]. An example of alternative Poincare plot-based classification considering both temporal and voltage differences was developed by Mensing et al. [5]. A sophisticated beat morphology-based classification method was proposed by de Chazal et al. [6]. This algorithm first detects ECG waves and then separately considers respective time and value features. Time-frequency domain was proposed as another feature space for ECG beats classification [7, 8]. An efficient feature space for beat classification also results from relationships of neighbouring coefficients [9] in so called cone of influence [10]. Considering a multilead ECG record opens the opportunity for inclusion inter-channel relations (e.g. angular) to the feature space. An aggregate of temporal, morphological and time-frequency features was proposed by Llamedo-Soria et al. [11, 12, 13]. Another approach worth to be mentioned here is based on syntactic model of the ECG [14], where selected parameters of best fitted figures are used as classification features.

While numerous approaches reported in the literature aimed at discovering a universal most-discriminative parameter (or set of parameters), we propose to use genetic algorithm to select a very small patient-dependent feature set of high discrimination capabilities.

2. Materials and methods

2.1. Genetic algorithm with aggressive mutation

The genetic algorithm with aggressive mutation was originally presented in [15] for determining of most distinctive EEG features for human-machine interface. In [16] it was compared with other algorithms in this domain showing unbeatable results. Similarly to the brain pattern recognition, identification of most distinctive ECG features leads to faster and more robust algorithms for beats classification. We applied the algorithm as originally proposed, but extended its applicability with individual learning step allowing for personalization of classification procedure.

Genetic algorithms are group of heuristic methods for solving optimization problems originating from genetic sciences, but currently adopted for other research fields, including feature selection process. In this application each individual encodes one subset of features and the algorithm aims to produce new individuals, represented by features of higher discrimination capabilities. A classifier has to be implemented for assessment of discriminative power of each individual. A common approach uses the classic genetic algorithm [17], where the genes of an individual represent the inclusion of corresponding feature of complete feature space to the considered feature subset. Unfortunately, the classic evolution controlled by classification accuracy favors individuals containing more features, thus reduction of feature set cannot be expected. The algorithm applied here proposes alternative, integer encoding of individuals in predefined small number of genes and modification of mutation scheme yielding a larger mother population.

In the aggressive mutation the mother population of M individuals is transformed into new population, but in the opposite to the classic approach, each parent of N genes has N children resulting from mutating its another gene. The size of the new population is thus $M \times N$. Another M individuals are created during the crossover operation. All these individuals are then compared using classification accuracy as a fitness function, and M individuals with best results are retained as mother population for next iteration. Since all individuals from the mother population are considered in the evaluation step, the value of classification accuracy in next mother population doesn't decrease.

2.2. Learning and classification procedure

The classification uses subject-optimized features of the ECG, thus it requires a learning phase in order to determine the feature set based on the present record. The feature set is expected to have as few elements as possible for a reliable distinction of all heart beats in the record. The learning is performed once and feature selection remains valid until the changes in subject status or recording conditions cause classification failures.

The learning is performed off-line and preceded with building of learning and testing sets containing exemplary heart beats from the record. At first, several individual ECG strips are randomly selected in a 24h record and classified with a conventional kNN-based algorithm using voltage difference as a feature. The resulting classes are memorized as a patient-specific locally-optimal groups. If the ECG record is stable, the heart beats from subsequent strips fall in the same classes and no new classes are created. Consequently, the initial locally-optimal groups are also optimal for the whole record. Otherwise, when new heart beats don't fit for existing classes, the initial and the problematic strips are concatenated and the classification algorithm is run again as many times as necessary to determine a globally-optimal groups. At this stage classification results are optionally presented to the human operator for confirmation of correctness. It is noteworthy that only few hundreds of beats from the whole record are classified in this manner. Finally, 35% of the classified heartbeats are selected as a learning set, while the remaining 65% belong to testing set for next procedure optimizing the feature set (fig. 1).

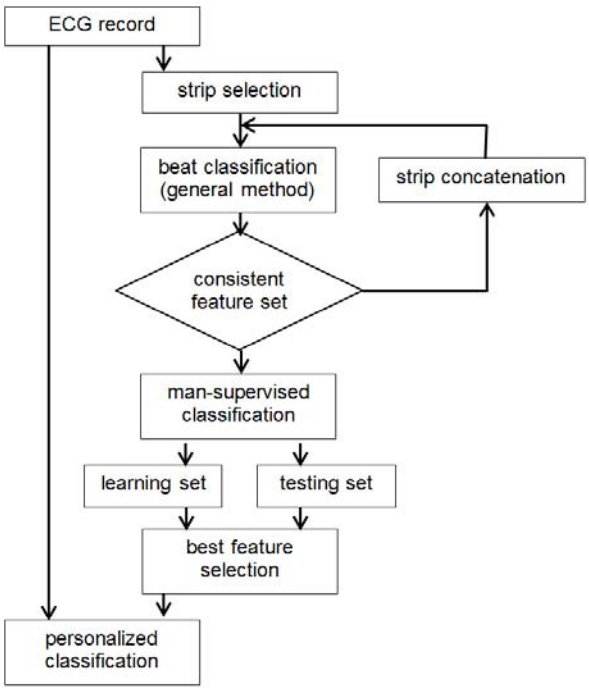

Figure 1 Diagram flow of feature set optimization for personalized heart beats classification.

Optimization of feature set calculates all considered features for all heart beats in the learning set and sorts them by descending discrimination power. Consequently,

first most discriminative features constitute a minimal set allowing for reliable classification of all beats. If the classification fails or if the distance between classes doesn't guarantee the expected reliability, next feature from the ordered list is taken into account. The optimization is performed by the genetic algorithm with aggressive mutation presented in 2.1, selecting given number of most distinctive features (3-5 depending on the experiment setup). These features were then used to classify the beats from testing set. If the classification accuracy measured on the heart beats from the testing set meet the quality criteria, the resulting feature set is used for classification of the whole record.

At this stage the resulting feature set may be considered as optimal for a given patient and given recording conditions. It may be used for classification of further ECG without the need of re-learning as long as the relationship between the signal source and the measurement system remains unchanged.

3. Tests and results

3.1. Testing conditions

The proposed method was tested with the MIT-BIH Arrhythmia database featuring reference beat types for each record. We applied a general heart beat detection procedure [18] and implemented algorithms calculating 57 different parameters of the beat:

- 4 derived from the QRS shape,
- 7 based on the signal derivatives (acceleration),
- 3 based on the wave area,
- 12 calculated from interval lengths,
- 5 derived from angles between two leads,
- 4 derived from Poincare plots of waves,
- 16 calculated from time-frequency representation,
- 6 calculated as parameters of best fitted dual channel syntactic model.

In two independent approaches we used common kNN [19] and linear SVM [20] as classification methods. Man-supervised recruitment of beats for the learning and testing sets was based on randomly selected non-overlapping 15 strips each of 10s length. Both sets (containing ca. 180 beats altogether) were carefully inspected for misclassified beats. The optimization of feature set was performed by the aggressive mutation algorithm with 100 iterations run for 60 individuals form the learning set. We assumed that an optimized feature set contains 3, 4 or 5 elements, so the initial feature space was reduced 11-19 times.

Reference classification method R1 was based on principal feature analysis i.e. best features were selected in the whole MIT-BIH Arrhythmia database by their pairwise minimum cross correlation [21]. In order to reveal the benefit of subject-oriented optimization of the feature selection, a second reference R2 was made of the best features selected by the same GAAM algorithm, but with the universal learning set composed of heart beats originating from various records of the MIT-BIH Arrhythmia database.

.

3.2. Test results

Comparing to the classification based on reference minimum correlation-based selection of features we observe significant reduction of misclassified beats ratio (for SVM from 2.73% to 0.69% in average). Test results are presented in table 1 as the ratio of the total misclassified beats to the total count of beats.

Table 1. Percentage of misclassified heart beats

classifier	no. of features	reference R1	globally optimal	personally optimal
kNN	3	3.11	2.24	1.15
	4	2.84	2.08	1.03
	5	2.57	1.89	0.98
SVM	3	2.73	1.55	0.69
	4	2.29	1.31	0.61
	5	2.01	1.18	0.58

Another interesting result was the average ranking of features by their discrimination power. As a reference we used the globally optimal feature set selected by the same GAAM algorithm. The RR_{n+1}/RR_n ratio was found the most discriminative parameter with the global optimization approach. Table 2 lists the positions taken by this parameter in the ranking of discriminative power resulting from personal optimization.

Table 2. Positions taken by the RR_{n+1}/RR_n ratio in the ranking of discriminative power resulting from personal optimization (the total no. of records equals 44).

classifier	no. of features	position				
		1-st	2-nd	3-rd	4-th	5-th
kNN	3	29	7	3		
	4	29	6	2	2	
	5	28	6	3	2	1
SVM	3	32	10	2		
	4	31	9	3	1	
	5	30	8	3	2	1

4. Discussion

Reduction of the feature set size reveals that even using a weaker classifier (kNN versus SVM) the 3-elements optimized set (tab. 1 col. 4 row 1) is still more distinctive than the 5-elements universal reference (tab. 1

col. 3 row 6). Since the optimization stage is performed once (unless the patient has significant progress of cardiac disease), the algorithms calculating the optimal features may accompany the personal data or may be embedded in reprogrammable personal ECG recorder.

The originality of presented approach consists in an individual, subject-dependent selection of most distinctive ECG features, instead of using of a standardized feature set. For repeatability of the results, these features have to be accommodated by classification algorithm e.g. programmed into a personal ECG recorder or supported by general-purpose interpretive equipment.

Acknowledgements

This Scientific work is supported by the AGH University of Science and Technology in year 2014 as a research project No. 11.11.120.612.

References

[1] Moody GB. The MIT-BIH arrhythmia database CD-ROM. Third Ed., Harvard-MIT Division of Health Sciences and Technology, 1997

[2] O'Dwyer M, de Chazal P, Reilly RI. Beat Classification for use in arrhythmia analysis. Computers in Cardiology 2000;27:395-398.

[3] Chang KC, Lee RG, Wen C, Yeh MF. Comparison of similarity measures for clustering electrocardiogram complexes. Computers in Cardiology 2005;32:759−762

[4] Lemay M, Jacquemet V, Forclaz A, Vesin JM, Kappenberger L. Spatiotemporal QRST cancellation method using separate QRS and T-Waves templates. Computers in Cardiology 2005;32:611−614.

[5] Mensing S, Bystricky W, Safer A. Identifying and measuring representative QT intervals in predominantly non-normal ECGs. Computers in Cardiology 2006;33:361−364.

[6] de Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering2004;51:1196-1206.

[7] Llamedo-Soria M, Martínez JP. An ECG Classification model based on multilead wavelet transform features Computers in Cardiology 2007;34:105−108.

[8] Rodrıguez-Sotelo JL, Cuesta-Frau D, Castellanos-Domınguez G. An improved method for unsupervised analysis of ECG beats based on WT features and J-means clustering. Computers in Cardiology 2007;34:581−584.

[9] Vansteenkiste E, Houben R, Pizurica A, Philips W. Classifying electrocardiogram peaks using new wavelet domain features. Computers in Cardiology 2008;35:853−856.

[10] Mallat S, Zhong S. Characterization of signals from multiscale edges. IEEE Trans on Pattern Anal and Machine Intel 1992;14:710–732.

[11] Llamedo -Soria M, Martínez JP. Analysis of multidoma in features for ECG classification. Computers in Cardiology 2009;36:561−564.

[12] Llamedo M, Khawaja A, Martínez JP. Analysis of 12-lead classification models for ECG classification. Computing in Cardiology 2010;37:673−676.

[13] Llamedo M, Martínez J. Heartbeat classification using feature selection driven by database generalization criteria. IEEE Transactions on Biomedical Engineering 2011;58: 616 - 625.

[14] Jokić S, Krčo S, Delić V, Sakač D, Lukić Z, Loncar-Turukalo T. An efficient approach for heartbeat classification. Computing in Cardiology 2010;37:991−994.

[15] Rejer I. Genetic algorithms in EEG feature selection for the classification of movements of the left and right hand. In Proc. CORES 2013: 579-589. doi: 10.1007/978-3-319-00969-8_57

[16] Rejer I, Genetic algorithm with aggressive mutation for feature selection in BCI feature space pattern analysis and applications (in print).

[17] Holland JH, Reitman JS, Cognitive systems based on adaptive algorithms. ACM SIGART Bulletin 63, June 1977:49

[18] Martínez A, Alcaraz R, Rieta JJ. Application of the phasor transform for automatic delineation of single-lead ECG fiducial points. Physiol Meas 2010;31:1467–1485.

[19] MacQueen, J. Some methods for classification and analysis of multivariate observations. [in:] L. M. Le Cam & J. Neyman [eds.] Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, Vol. 1, University of California Press, Berkeley 1967: 281-297.

[20] Vapnik VN. The nature of statistical learning theory. Springer-Verlag New York, Inc., New York, NY, USA. 1995

[21] Augustyniak P. The use of shape factors for heart beats classification in Holter recordings. Computers in Medicine Zakopane 2-6. 05. 1997:47-52.

Address for correspondence.

Piotr Augustyniak
AGH University of Science and Technology
30 Mickiewicz Ave., 30-059 Kraków, Poland.
august@agh.edu.pl