Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 44 (2023), No. 4, pp. 607–623
DOI : 10.47974/JIOS-1128

Non-stationary wavelet for ECG signal classification

Abdelmalik Boussaad ‡
Laboratory of Applied Mathematics
University of Biskra
145 RP
07000 Biskra
Algeria

Khaled Melkemi †
EDPA Laboratory
University of Batna 2
53 Constantine Road
Batna 05000
Algeria

Farid Melgani §
Department of Information Engineering and Computer Science
University of Trento
Via Sommarive
9 - 38123 Povo Trento
Italy

Zouhir Mokhtari *
Laboratory of Applied Mathematics
University of Biskra
145 RP
07000 Biskra
Algeria

‡ *E-mail:* a.boussaad2001@gmail.com
† *E-mail:* kmelkemi@yahoo.fr
§ *E-mail:* melgani@disi.unitn.it
* *E-mail:* z.mokhtari@univ-biskra.dz (Corresponding Author)

Abstract

Wavelet analysis has shown to be an interesting tool for representing ECG signals for classification. In this paper, we present a new ECG signal representation based on the notion of non-stationary wavelets. The main difference with the construction of standard wavelets is that the multiresolution spaces are generated by scale-dependent functions in order to achieve increased flexibility and sparseness. In order to customize the non-stationary wavelet to the given ECG classification task, we resort to the fireworks optimization algorithm, thus making the proposed method general and not constrained by the choice of a particular classifier. The proposed method is validated on AAMI classes of the well-known MIT data set. Results compared to standard stationary wavelets show a significant boost in accuracy.

Subject Classification: *[2010] 65T60, 68T05, 74P10, 42C40.*

Keywords: *Discrete wavelet transform (DWT), ECG signal classification, Extreme learning machine (ELM), Fireworks algorithm for optimization (FA), Non-stationary wavelets.*

1. Introduction

The automatic classification of electrocardiographic (ECG) signals has attracted a great attention of many researchers in the biomedical engineering domain since numerous years. This interest is mainly explained by the fact that ECG signal is an important source which provides cardiologists with valuable information about the functional aspects of the heart and the cardiovascular system. Accordingly, many researches have been focused on developing models for ECG classification in order to design automatic and accurate cardiac arrhythmias diagnosis tools.

In the context of ECG signal classification, several interesting methods have been proposed, like those in [1, 2]. On the other hand, the literature reports also a great interest in exploiting wavelets as a feature extraction method [3, 4, 5, 6] in different classification areas. Particulary, concerning ECG classification, *Ince et. al.* in [3] presented a feature extraction technique which employs the translation-invariant dyadic wavelet transform (TIDWT) in order to effectively extract the morphological information from ECG data. Further, a wavelet filter bank design for ECG signal classification method is proposed in [4]. This method exploits the polyphase representation of the discrete wavelet transform (DWT), that allows generating a wavelet filter bank from a set of angular parameters and thus formulating the wavelet design problem for ECG signal classification as a problem of estimating these parameters. From these works, it appears clearly that, for classifying ECG heartbeat patterns the representation of the signal is very important to extract features. Fourier transform only provides the spectral components of the signal, not their temporal relationships [7]. However wavelets can provide a time-frequency representation of the signal [8]. Most of the aforementioned algorithms which integrate wavelets in the classification process have proved their effectiveness in the area of ECG classification. In this context, our paper proposes a method for generating non-stationary wavelets proposed by [9] for representing ECG signals. Wavelets generation is guided by the classification process performance in term of accuracy. This approach is based on the choice of the best parameter vector of the exponential polynomials to be reproduced by the scaling functions [9], that allows to construct a non-stationary wavelet with fixed scaling filter length in each decomposition level. The estimation of these parameters is carried out through the maximization of the classifier accuracy. A fireworks algorithm for optimization (FA) [10] is adopted to optimize the fitness function. For experimental validation, we resorted to the extreme learning machine (ELM) classifier for its effectiveness.

The rest of the paper is organized as follows: Section 2 presents the basic definitions with different steps for constructing non-stationary wavelets. Section 3 presents the fireworks algorithm for optimization (FA). Section 4 describes the proposed non-stationary wavelet design method. The experimental part is reported in Section 5. Finally, conclusions are drawn in Section 6.

2. Non-stationary wavelets

2.1 *General concepts*

Non-stationary wavelets is a generalization of the Daubechies wavelets family introduced recently in [9]. The main difference with the standard construction of Daubechies is that the multiresolution analysis spaces are generated by scale-dependent functions. This implies that Mallat's fast wavelet transform algorithm can be reformulated using scale-dependent filter banks. In the following, we will present the basic definitions and the most important steps for the construction of non-stationary wavelets proposed by [9].

Definition 1 : (Non-stationary multiresolution analysis)

Given $(\varphi_j)_{j\in\mathbb{Z}}$ a sequence of scaling functions. The subspaces $(V_j)_{j\in\mathbb{Z}} \subset L^2(\mathbb{R})$ such that:

$$V_j = Span\left\{ \frac{1}{\sqrt{2^j}} \varphi_j(2^{-j}x - k) / k \in \mathbb{Z} \right\} \tag{1}$$

define a non-stationary multiresolution analysis of $L^2(\mathbb{R})$ if, for any $j \in \mathbb{Z}$

1. $\left\{ \frac{1}{\sqrt{2^j}} \varphi_j(2^{-j}x - k), k \in \mathbb{Z} \right\}$ is an orthonormal basis of V_j

2. $$V_{j+1} \subset V_j \tag{2}$$

3. $$\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R}) \quad \text{and} \quad \bigcap_{j\in\mathbb{Z}} V_j = \{0\} \tag{3}$$

The embedding (2) implies that for any $j \in \mathbb{Z}$ there exists a set of scaling filters $\{h_k^j\}_{k\in\mathbb{Z}}$ such that:

$$\varphi_{j+1}(2^{-(j+1)}x) = \sum_{k\in\mathbb{Z}} h_k^j \varphi_j(2^{-j}x - k) \tag{4}$$

In the non-stationary multiresolution analysis case, the scaling filters and the scaling equations (4) are scale-dependent. Like in the standard case for each level *j*, the orthogonal complement of V_j in V_{j-1} is W_j such that:

$$W_j = Span\left\{ \frac{1}{\sqrt{2^j}} \psi_j(2^{-j}x - k) / k \in \mathbb{Z} \right\} \tag{5}$$

Where,

$$\psi_{j+1}(x) = \sum_{k\in\mathbb{Z}} g_k^j \varphi_j(2x - k) \quad \text{and} \quad g_k^j = (-1)^k \overline{h_k^j} \tag{6}$$

The sets of coefficients $\{h_k^j\}_{k\in\mathbb{Z}}$ and $\{g_k^j\}_{k\in\mathbb{Z}}$ form a non-stationary filter bank.

$H_j(z) = \sum_{k\in\mathbb{Z}} h_k^j z^k$ and $G_j(z) = \sum_{k\in\mathbb{Z}} g_k^j z^k$ are their z-transform respectively. ψ_j is called the non-stationary wavelet. The *j* level DWT decomposition of a signal $x(t)$ in the case of non-stationary wavelet is shown in Figure 1.

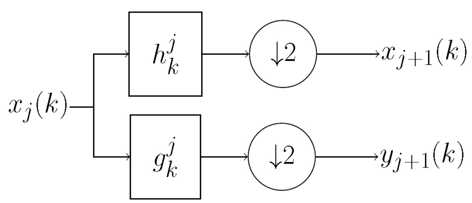

Figure 1

Discrete wavelet transform at level j in the case of non-stationary wavelet

Definition 2 : (Reproduction of exponential polynomials)

Given a vector $\vec{\alpha} \in \mathbb{C}^n$ with $\alpha_{(m)}$ its distinct components and n_m their respective multiplicities, the non-stationary multiresolution $(V_j)_{\in\mathbb{Z}}$ reproduces the exponential polynomials corresponding to $\vec{\alpha}$ if and only if, for an $P(x)e^{\alpha_{(m)}x}$, where $P(x)$ is a polynomial of $deg < n_m$ and at any scale *j*, there exists a sequence $\{b_k^j\}_{k\in\mathbb{Z}}$ such that:

$$P(x)e^{\alpha_{(m)}x} = \sum_{k\in\mathbb{Z}} b_k^j \varphi_j(2^{-j}x - k) \tag{7}$$

The method proposed in [9] is a technique to design non-stationary multiresolution spaces under the following constraints on the scaling functions :

1. Reproduce a given set of exponential polynomials corresponding to $\vec{\alpha}$ which does not depend on the scale *j*.
2. The scaling functions φ_j are real-valued and compactly supported.
3. $\langle \varphi_j(x), \varphi_j(x-k)\rangle = \delta(k)$

$$\text{Where} \qquad \delta(k) = \begin{cases} 1 & \text{if} \quad k = 0 \\ 0 & \quad\quad \text{Otherwise} \end{cases}$$

In [9] authors proved that the previous constraints are equivalent to:

1. Each filter H_j has a zero of order n_m at $z = -e^{2^j \alpha_{(m)}}$ for all *m*, this means that H_j must be divisible by $B_{2^j \bar{\alpha}}$ where,
$$B_{2^j \bar{\alpha}}(z) = \prod_{k=1}^{n} (1 + e^{2^j \alpha_{(k)}} z^{-1}) \tag{8}$$
2. The elements of $\vec{\alpha}$ should be real or appear in complex conjugate pairs with the same multiplicity and $H_j(z)$ is a finite impulse response filter (FIR).
3. For any $j \in \mathbb{Z}$:
$$H_j(z)H_j(z^{-1}) + H_j(-z)H_j(-z^{-1}) = 4 \tag{9}$$

From the imposed previous constraints, the design problem can be formulated as follows.

For a given vector $\vec{\alpha}$ that satisfies the constraint (2) and a level *j*, find the shortest FIR $Q_j(z)$ such that:

$$\begin{cases} H_j(z)H_j(z^{-1}) + H_j(-z)H_j(-z^{-1}) = 4 \\ H_j(z) = B_{2^j \vec{\alpha}}(z)Q_j(z) \end{cases} \tag{10}$$

The problem in (10) is equivalent to:

$$C_j(Z)D_j(Z) + C_j(-Z)D_j(-Z) = 2 \tag{11}$$

where $Z = \frac{z+z^{-1}}{2}$, $C_j(Z) = c_j B_{2^j \vec{\alpha}}(z) B_{2^j \vec{\alpha}}(z^{-1})$, $2c_j D_j(z) = Q_j(z)Q_j(z^{-1})$

and $c_j = \prod_{k=1}^{n} e^{-2^j \alpha_k}$. The equation (11) is called the Bezout equation.

An algorithm for solving the design problem (10) was introduced in [9]. It is based on the following steps:

Let $\vec{\alpha} \in \mathbb{C}^n$ a vector which satisfies the constraint (2) and j is a level index

Step 1: If the concatenated vector $\vec{\delta} = [\vec{\alpha}, -\vec{\alpha}]$ cannot comprise components δ and δ' such that:

$$2^j(\delta - \delta') = i(2k+1)\pi, \quad \text{for} \quad k \in \mathbb{Z} \tag{12}$$

solve for $D_j(Z)$ equation (11).

Step 2: If

$$D_j(Z) \geq 0 \quad \text{over } [-1,1] \tag{13}$$

perform the spectral factorization $2c_j D_j(Z) = Q_j(Z)Q_j(Z^{-1})$ to obtain $Q_j(Z)$.

Thus, the scaling filter $H_j(Z)$ is calculated by $H_j(z) = B_{2^j \vec{\alpha}}(z)Q_j(z)$, the sequences $\{h_k^j\}_{k\in\mathbb{Z}}$, $\{g_k^j\}_{k\in\mathbb{Z}}$ and ψ are obtained by the same process in the case of standard wavelets for each level j.

For such purpose, in our construction we will develop a technique based on fireworks algorithm to select a vector $\vec{\alpha} \in \mathbb{C}^n$ in the search space which holds the conditions (12) and (13) for every level j and reproduce a scaling filter with length $4n$ for best ECG signal classification.

3. Fireworks Algorithm for Optimization

Fireworks algorithm (FA) is an evolutionary computation algorithm like as genetic algorithms (GAs) [11] or particle swarm optimization (PSO) [12], which is inspired from swarm intelligence. It was recently proposed by [10] for solving complex design optimization problems in different applications, such as electrical and mechanical engineering [13, 14].

As explained in [10], the main concepts of the FA algorithm are given below:

Let us consider the following optimization problem:

$$\min_{x\in\mathbb{R}^d} f(x) \qquad \text{such that}: x_{\min} \leq x \leq x_{\max}$$

where: $x = (x^1, x^2, ..., x^d)$ represents a location in the potential space (or a firework),

$x_{\min} = (x^1_{\min}, x^2_{\min}, ..., x^d_{\min})$, $x_{\max} = (x^1_{\max}, x^2_{\max}, ..., x^d_{\max})$ denote the bounds of the potential space and $f(x)$ is an objective function.

The constraint $x_{\min} \leq x \leq x_{\max} \Leftrightarrow (x^1_{\min} \leq x^1 \leq x^1_{\max}) \wedge ... \wedge (x^d_{\min} \leq x^d \leq x^d_{\max})$.

Generally, the FA algorithm consists of three steps, explosion step (generating explosion sparks), Gaussian mutation and selection strategy.

3.1 *Explosion step*

In this part of the FA algorithm, the amplitude of explosion A_i and the number of sparks S_i for each firework x_i are defined by:

$$A_i = \hat{A}\frac{f(x_i) - y_{\min} + \zeta}{\sum_{i=1}^{n}(f(x_i) - y_{\min}) + \zeta} \tag{14}$$

$$S_i = m\frac{y_{\max} - f(x_i) + \zeta}{\sum_{i=1}^{n}(y_{\max} - f(x_i)) + \zeta} \tag{15}$$

where $y_{\min}$ and $y_{\max}$ are the minimum and maximum of the objective function *f* among the *n* fireworks, $\hat{A}$ is a real constant represents the sum of all $A_i / i = 1, n$ and *m* is the total number of sparks, while ζ is a small constant used to prevent the zero-division. The approximation and the limitation of number of sparks are defined as follows:

$$\hat{S}_i = \begin{cases} \text{round}(\alpha.m) & \text{if} \quad S_i < \alpha.m \\ \text{round}(\beta.m) & \text{if} \quad S_i > \beta.m \\ \text{round}(S_i) & \text{Otherwise} \end{cases} \tag{16}$$

where α and β are constants such that $\alpha < \beta < 1$, and round() is the rounding function. Algorithm 1 describes the manner of generating sparks in this step of the FA algorithm.

Algorithm 1 Explosion step

1: Initialize location of "explosion spark $\hat{x}_j$" $\hat{x}_j = x_i$.
2: Set $z^k = \text{round}(\text{rand}(0,1)), k = 1, 2, \ldots, d$
3: **for** each dimension of $\hat{x}_j$ where $z^k == 1$ **do**
4: Calculate $h^k = A_i * \text{rand}(-1,1)$
5: $\hat{x}_j^k = \hat{x}_j^k + h^k$
6: **if** $\hat{x}_j^k < x_{\min}^k$ or $x_j^k > x_{\max}^k$ **then**
7: randomly map $\hat{x}_j^k$ to the search space: $\hat{x}_j^k = x_{\min}^k$ + rand(0,1) $(x_{\max}^k - x_{\min}^k)$ // rand(0,1) denotes a random number from uniform distribution in $[0,1]$ //.
8: **end if**
9: **end for**

3.2 *Gaussian mutation step*

As shown in[10], to ensure the diversity of sparks, FA algorithm uses another kind of sparks, which are called Gaussian sparks. Algorithm 2 shows the manner of generating this kind of sparks.

Algorithm 2 Gaussian mutation step

1: Initialize location of "a specific spark $\hat{x}_j$" $\hat{x}_j = x_i$.
2: Set $z^k = \text{round}(\text{rand}(0,1)), k = 1, 2, \ldots, d$
3: Calculate the coefficient $g = \mathcal{N}(1,1)$
4: **for** each dimension of $\hat{x}_j$ where $z^k == 1$ **do**
5: $\hat{x}_j^k = \hat{x}_j^k * g$
6: **if** $\hat{x}_j^k < x_{\min}^k$ or $x_j^k > x_{\max}^k$ **then**
7: randomly map $\hat{x}_j^k$ to the search space: $\hat{x}_j^k = x_{\min}^k$ + rand(0,1) $(x_{\max}^k - x_{\min}^k)$
8: **end if**
9: **end for**

3.3 *Selection strategy step*

At each iteration FA keeps the best location for the next explosion generation, among all the current sparks and fireworks, then $n-1$ locations are selected with some probabilities proportional to their distances to other location. The selection probability is defined by the following equations:

$$R(x_i) = \sum_{j \in k} d(x_i, x_j) = \sum_{j \in k} \left\| x_i - x_j \right\| \tag{17}$$

$$p(x_i) = \frac{R(x_i)}{\sum_{j \in K} R(x_j)} \tag{18}$$

Finally, we summarize the FA algorithm in the following pseudo-code

Algorithm 3 FA algorithm

Randomly select n locations for fireworks
while stop creteria=false **do**
for each firework **do**
Explosion step
end for
Randomly select $\hat{m}$ fireworks for generating Gaussian sparks
for $k = 1 : \hat{m}$ **do** Gaussian mutation step
end for
Calculate $f(x_i)$ for all fireworks and sparks
Selection strategy step
end while

4. Proposed method

In this section, we describe the proposed non-stationary wavelet design method for the classification of ECG signals. As mentioned in the introduction, the literature reports many methods exploiting wavelets for analysing ECG signals. Most of these techniques use the Discrete Wavelet Transform as a tool for representing signals, which is equivalent to decompose the signal with Mallat's algorithm up to a desired level. In the case of standard wavelets, the scaling filters and the wavelet filters used in Mallat's algorithm are the same for each decomposition level. However, the decomposition of the ECG signal through a non-stationary filter bank, which means using Mallat's algorithm with different filters for each level, could offer a more customized and thus better representation in particular if it is task-driven. In this context, we propose a non-stationary wavelet design which reproduces a different filter in each level for Mallat's decomposition algorithm. The design process is driven by the ECG classification accuracy using fireworks algorithm for optimization. As shown in Figure 2, the vectors $\{\alpha_i\}_{i=1,n}$ represents locations of fireworks.

Concerning the fitness function, we use the classification accuracy of the ELM classifier. It is noteworthy that the proposed method is general and thus not constrained by the use of a specific classifier (in our experiments the ELM classifier will be adopted, but any other classifier could be exploited as well). For more details concerning ELM classifier, see [15, 16]

The fitness function is an estimation of the accuracy which drives the FA algorithm to select the best wavelet candidate. During the training phase classifier hyperparameters (in the case of ELM, regularization and kernel parameters) are chosen in the wavelet domain by a standard cross-validation procedure. As illustrated in Figure 2, the main steps of our proposed method are described in the following:

Algorithm 4 Proposed method

Set the length of decomposition filters $= 4n$, and the decomposition level $= d$.
Initialize the FA algorithm with n random population vectors which represent locations of fireworks $\{v_k = (\alpha_1^k, \alpha_2^k, \dots, \alpha_n^k) / k = 1, \dots, n\}$.
Generate sparks for each firework according to Algorithm 1.
Generate Gaussian sparks according to Algorithm 2.
for each spark and firework location $(\alpha_1^k, \alpha_2^k, \dots, \alpha_n^k)$ **do**.
for each level j **do**.
Set the vector $\vec{\alpha}^j = 2^j(-i\alpha_1^k, i\alpha_1^k, -i\alpha_2^k, i\alpha_2^k, \dots, -i\alpha_n^k, i\alpha_n^k)$.
Find the shortest Bezout solution $D_0^{\vec{\alpha}^j}$.
end for.
end for.
for each spark and firework location $(\alpha_1^k, \alpha_2^k, \dots, \alpha_n^k)$ **do**.
if $D_0^{\vec{\alpha}^j}$ exists and positive over $[-1,1]$ at every j **then**.
for each level j **do**.
Generate the filter bank $\{h_k^j, g_k^j\}$ of level j.
end for.
Apply DWT to each ECG training beat up to level d.
Train an classifier (e.g ELM) by feeding it with the generated wavelet features and compute its cross-validation.
Evaluate the fitness function $F(\alpha_1^k, \alpha_2^k, \dots, \alpha_n^k)$.
else.
Set fitness function $F(\alpha_1^k, \alpha_2^k, \dots, \alpha_n^k) = 0$.
end if.
end for.

23: **if** FA converge **then.**
24: Select the best non-stationnary filter bank according to the best location.
25: **else.**
26: Select the best locations according to Eq 18.
27: Generate a new population and go to 03.
28: **end if.**

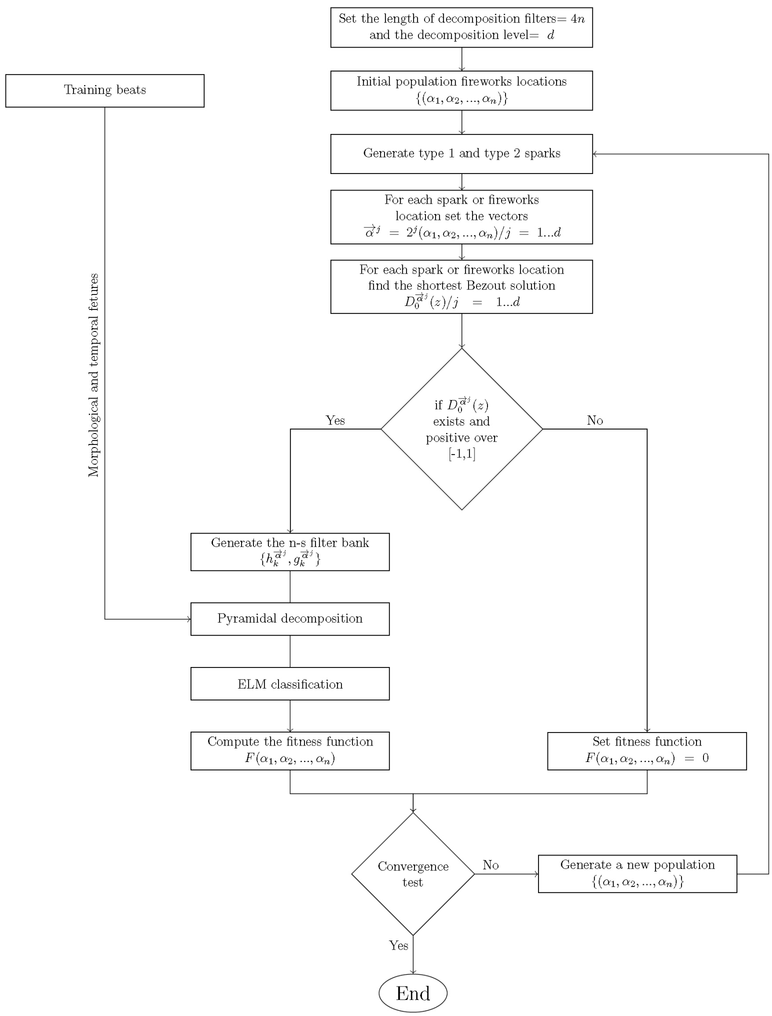

Figure 2

Block diagram of the proposed non-stationary wavelet design guided by the ECG classification process

5. Experiment setup and performance evaluation

In our proposed algorithm we are interested in distinguishing the following four AAMI (Association for the Advancement of Medical Instrumentation) classes of the MIT database, namely the (*N*) normal class, (*S*) supraventricular class, (*V*) ventricular class and (*F*) fusion class. The initial training set is composed of 22 records, namely DS1=101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230 of MIT-BIH, the others records of this database are left for test which are grouped in the following set DS2=100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 201, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234. In this context, for generating the training set from DS1 we follow the same method proposed by [17], which is based on clustering the global training set with the k-means clustering algorithm. In particular, for each AAMI class, the k-means algorithm is run to split it into Nclust clusters. Then the closest samples to the centroid of each cluster are selected to form a training set with a reduced size. This choice offers a good compromise between computation and classification accuracy. In our case, we perform a clustering in the training set DS1 according to cluster numbers reported in Table1.

Table 1

Number of clusters for each class

Class	N	S	V	F
Number of clusters	2000	500	1000	200

Accordingly, the training set is formed of 3700 beats, while the test set contains all the DS2 beats, namely 49664 beats.

The performance of our proposed algorithm to discriminate the different classes of AAMI was calculated using various measures, which are :

1. The overall accuracy (*OA*), which represents the percentage of correctly classified beats among all the beats considered (independently with respect to the classes they belong to).
2. The sensitivity $S_e = \frac{T_P}{T_P + F_N}$, it is the ratio between the number of correctly detected beats concerning a class and the total number of beats belong to this class.

3. The positive predictivity, $P_p = \frac{T_P}{T_P + F_P}$, namely the ratio between the number of correctly detected beats concerning a class and the total number of detected beats of this class .

The above notations are as follows: TP = true positive, FN = false negative and FP = false positive.

5.1 *Experimental results and discussion*

Our proposed non-stationary wavelet (Nsw) is compared to the reference Daubechies wavelet (db2 and db4 of scaling filters of length 4 and 8, respectively) of same length giving rise to two non-stationary wavelets termed in the following as Nsw2 and Nsw4 and by considering different decomposition levels. The achieved results are summarized in Tables (2,3) and (4).

Table 2

Classification results obtained by the ELM classifier on the test set DS2 using original signal (Orig) and Db2, Db4 and proposed Nsw2, Nsw4 wavelets features in level 1

	N		S		V		F		OA
	Se	Pp	Se	Pp	Se	Pp	Se	Pp	
Orig	73.24	92.88	60.03	87.35	75.49	98.35	57.05	87.30	72.90
Db2	74.43	90.01	55.85	89.53	70.46	98.05	59.28	86.13	76.46
Nsw2	**81.50**	**90.03**	**61.57**	**94.18**	**79.05**	**99.03**	**60.08**	**93.24**	**80.86**
Db4	75.20	91.04	54.30	89.77	71.54	98.44	60.13	86.85	77.12
Nsw4	**81.56**	**91.50**	**61.60**	**95.06**	**79.40**	**99.01**	**61.76**	**94.33**	**81.43**

Table 3

Classification results obtained by the ELM classifier on the test set DS2 using original signal (Orig) and Db2, Db4 and proposed Nsw2, Nsw4 wavelets features in level 2

	N		S		V		F		OA
	Se	Pp	Se	Pp	Se	Pp	Se	Pp	
Orig	73.24	92.88	60.03	87.35	75.49	98.35	57.05	87.30	72.90
Db2	77.35	98.87	56.14	90.27	70.36	98.01	60.15	86.07	76.58
Nsw2	82.77	**98.84**	67.43	94.70	78.12	98.25	61.20	93.42	82.10
Db4	**77.45**	98.90	**55.99**	**91.90**	**74.20**	**98.05**	**62.05**	**86.90**	**77.51**
Nsw4	**83.40**	98.87	**67.73**	**95.83**	**80.07**	**99.05**	**64.23**	**95.04**	**84.47**

Table 4

Classification results obtained by the ELM classifier on the test set DS2 using original signal (Orig) and Db2, Db4 and proposed Nsw2, Nsw4 wavelets features in level 3

	N		S		V		F		OA
	Se	Pp	Se	Pp	Se	Pp	Se	Pp	
Orig	73.24	92.88	60.03	87.35	75.49	98.35	57.05	87.30	72.90
Db2	77.60	98.85	54.55	90.41	69.94	98.30	59.20	86.91	76.56
Nsw2	**86.11**	**98.60**	**70.32**	**94.11**	**80.35**	**98.81**	**62.01**	**92.85**	**86.60**
Db4	77.95	98.88	60.15	92.23	74.30	98.75	62.15	86.93	78.65
Nsw4	**88.04**	**98.64**	**70.56**	**96.46**	**85.11**	**99.35**	**65.67**	**96.20**	**87.84**

In general, as it can be observed from these quantitative results, our proposed nonstationary wavelets exhibit superior performance compared to the standard Daubechies wavelets Db2 and Db4 with the same length of scaling filter, this for all the four classes. Such a result reveals that wavelet customization is a valuable way to enhance the representation of the ECG signals and consequently the discrimination capability in the resulting representation space. This is made possible thanks to the non-stationary nature as well as the FA optimization of the filter bank, independently from the decomposition level. Indeed, for the three explored levels, Nsw outperforms Db for both filter sizes (length 4 and 8). It is noteworthy that the higher the decomposition level, the larger the gap of accuracy between Nsw and Db (from around 4% for level 1 to around 10% for level 3). This can be explained by the fact that Nsw leads to a sparser representation and as the decomposition level increases as sparseness becomes more marked, equipping the classifier with a potentially higher generalization capability. In greater detail, comparing for example Db2 and Nsw2 and looking at the third decomposition level (Table 4), the use of the original ECG signal to feed the classifier generates an overall accuracy of 72,9% due mainly to a problem of discrimination of the ’N’ class which is largely dominant in the MIT dataset. The jump from a temporal representation to a time-frequency one thanks to Db2 allows to mitigate this issue, since more than 4% of extra accuracy is captured for the ’N’ class impacting substantially on the OA (from 72,9% to 76.6%). Such positive gain is however paid by a significant decrease of sensitivity for the classes ’S’ and ’V’, which are less frequent in the dataset but not less important than the class ‘N’. The application of Nsw has allowed to bring a positive effect on all the classes in terms of sensitivity as well as predictivity, suggesting that the resulting feature space has strongly improved the separability between the class distributions and not just incurred in a partial class distribution shift. The same observations can be reported regarding Db4 with respect to Nsw4 (i.e., with 8 coefficients scaling filter length). The main limit of the proposed method lies in the heavy computational load required to generate the optimal non-stationary filter bank. Indeed, on a (Intel I5 processor of 1.70 GHz frequency with 8 GB of RAM memory), less than 20 minutes have been consumed to get the classifier ready for predicting the label of the test beats. However, since the training process is done once and offline, this kind of issue can actually be seen as not that critical in regard to the achieved accuracy gains.

6. Conclusion

In this paper, we have elaborated an original machine-learning inspired method to conceive an optimized wavelet-based classifier for ECG signals. Its main contribution is that standard wavelet feature extraction, though potentially interesting and effective, can still be further improved thanks to a relaxation of the stationarity constraint. Indeed, nonstationary wavelets permit to achieve a customized and sparser decomposition. In particular, the wavelet design process is driven by the ECG classification accuracy using fireworks algorithm for optimization. Though the ELM classifier was adopted in our experimental validation, it is noteworthy that the proposed method has a general nature and thus could be implemented with any classifier. The results show superior accuracies can be achieved, though with a substantial extra computational load we however consider acceptable. These promising results call for a more extensive experimental validation on other classifiers and other datasets (not necessarily limited to ECG signals), which however is out of the scope of the current manuscript. Another major future development regards the optimization process, which deserves to be simplified in order to make the method more attractive from a computational viewpoint.

References

[1] Farid Melgani and Yakoub Bazi. Classification of electrocardiogram signals with support vector machines and particle swarm optimization. *IEEE Transactions on Information Technology in Biomedicine*, 12(5) : 667-677 (2008).

[2] Argyro Kampouraki, George Manis and Christophoros Nikou. Heartbeat time series classification with support vector machines. *IEEE Transactions on Information Technology in Biomedicine*, 13(4) : 512-518 (2009).

[3] Turker Ince, Serkan Kiranyaz and Moncef Gabbouj. A generic and robust system for automated patient-specific classification of ecg signals. *IEEE Transactions on Biomedical Engineering,* 56(5) : 1415-1426 (2009).

[4] Abdelhamid Daamouche, Latifa Hamami, Naif Alajlan and Farid Melgani. A wavelet optimization approach for ecg signal classification. *Biomedical Signal Processing and Control,* 7(4) : 342-349 (2012).

[5] Behbod Keshavarzi, Behnam Zarpak, Seyed Hamid Haj Seyed Javadi and Gholam Reza Jandaghi. Adaptive steganography based on integer wavelet transform using block standard deviation. *Journal of Information and Optimization Sciences,* 40(7) : 1457-1473 (2019).

[6] MA Ansari, Rajat Mehrotra and Rajeev Agrawal. Detection and classification of brain tumor in MRI images using wavelet transform and support vector machine. *Journal of Interdisciplinary Mathematics,* 23(5) : 955-966 (2020).

[7] Kei-ichiro Minami, Hiroshi Nakajima and Takesshi Toyoshima. Real-time discrimination of ventricular tachyarrhythmia with fourier-transform neural network. *IEEE Transactions on Biomedical Engineering,* 46(2) : 179-185 (1999).

[8] L. Romero and L Serrano. Ecg frequency domain features extraction: A new characteristic for arrhythmias classification. Technical report, DTIC Document (2001).

[9] Cédric Vonesch, Thierry Blu and Michael Unser. Generalized daubechies wavelet families. *IEEE Transactions on Signal Processing,* 55(LIB-ARTICLE-2007-013) : 4415-4429 (2007).

[10] Ying Tan. Fireworks algorithm. *Heidelberg, Germany: Springer,* 10 : 978-3 (2015).

[11] DE Goldberg. Genetic algorithms in search, optimization and machine learning, addison-wesley, reading, ma, 1989. *K. Ohno, K. Esfarjani and Y Kawazoe: Computational Materi* (1992).

[12] James Kennedy, James F Kennedy, Russell C Eberhart and Yuhui Shi. *Swarm Intelligence*. Morgan Kaufmann (2001).

[13] A Mohamed Imran and M Kowsalya. A new power system reconfiguration scheme for power loss minimization and voltage profile enhancement using fireworks algorithm. *International Journal of Electrical Power & Energy Systems*, 62 : 312-322 (2014).

[14] Debkalpa Goswami and Shankar Chakraborty. Parametric optimization of ultrasonic machining process using gravitational search and fireworks algorithms. *Ain Shams Engineering Journal,* 6(1) : 315-331 (2015).

[15] Gao Huang, Guang-Bin Huang, Shiji Song and Keyou You. Trends in extreme learning machines: A review. *Neural Networks*, 61 : 32-48 (2015).

[16] Guang-Bin Huang, Hongming Zhou, Xiaojian Ding and Rui Zhang. Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, 42(2) : 513-529 (2011).

[17] Mohamad M Al Rahhal, Yakoub Bazi, Naif Alajlan, Salim Malek, Haikel Al-Hichri, Farid Melgani and Mansour A Al Zuair. Classification of aami heartbeat classes with an interactive elm ensemble learning approach. *Biomedical Signal Processing and Control*, 19 : 56-67 (2015).

Received May 2021

Revised February 2022