33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Feature Extraction in Time-Frequency Signal Analysis by means of Matched Wavelets as a Feature Generator.

Pawel S Kostka, Ewaryst J Tkacz, *Member, IEEE*

***Abstract*— The goal of presented work was to compare the usage of standard basic wavelet function like e.g. bio-orthogonal or dbn with the optimized wavelet created to the best match analyzing ECG signals in the context of P-wave and atrial fibrillation detection. A library of clinical expert evaluated typical atrial fibrillation evolutions was created as a database for optimal matched wavelet construction. Whole data set consisting of 40 cases with long term ECG recordings were divided into learning and verifying set for the multilayer perceptron neural network used as a classifier structure. Compared with other wavelet filters, the matched wavelet was able to improve classifier performance for a given ECG signals in terms of the Sensitivity and Specificity measures.**

I. Introduction

The subject of multiscale signal representation/analysis has been studied by applied mathematicians for a number of years. The works of Daubechies [1] and Mallat [2] evoked the interest of signal processing community in the theory of wavelet transforms. These papers established the connection between wavelet transforms and the theory of multirate filterbanks. Matching a wavelet to class of signals can improve the feature extraction stage in classifiers based on time-frequency signal decomposition as well as can increase the signal to noise ratio during de-nosing process. In signal detection applications, the decomposition of a signal in the presence of noise using a wavelet matched to the signal would produce a sharper and more discriminating peak in time-scale space as compared to standard non-matched wavelets [1]. Described in literature wavelet design techniques developed by Mallat and Zheng [3], and Chen and Donoho [4], build nonorthonormal wavelet bases from a library of existing wavelets in such a way that some error cost function is minimized. These techniques are constrained by the library of functions used and do not satisfy the need for optimal correlation in both scale and translation.

Proposed method as a modification of Chapa [5] approach bases on the matching the basic wavelet separately for magnitude and phase of the signal spectrum. Finding an optimized wavelet is a crucial stage of time-frequency ECG signal decomposition (feature extraction stage) of the atrial fibrillation episodes classifier structure. The goal of presented work was to compare the usage of standard basic wavelet function like e.g. bio-orthogonal or db*n* with the optimized wavelet created to the best match analyzing ECG signals in the context of P-wave and atrial fibrillation detection.

The application field of presented multi-domain feature extraction is the trial of detection of atrial fibrillation (AF), which is a supraventricular tachyarrhythmia characterized by uncoordinated atrial activation with consequent deterioration of atrial mechanical function. Last researches report AF as a result of a fractionated atrial electrical activity mainly due to the shortening of atrial refractory period, which allows multiple waves pass through the atrial mass. These changes ultimately reduce the inward calcium current, and this in turn reduces the action potential duration. If the action potential duration shortens, the refractory period shortens too, and the cell can be ready for reactivation earlier. On the electrocardiogram (ECG), AF is described by the replacement of consistent P waves by rapid oscillations or fibrillatory waves that vary in size, shape, and timing, associated with an irregular, frequently rapid ventricular response when atrio-ventricular conduction is intact. Because of disturbed haemodynamic, atrial fibrillation and atrial flutter are between of the most usual causes of thrombi-embolic events. So, development of methods supported its diagnosis seems to be still promising and meaningful research field.

II. Classifier structure with feature extraction stage based on matched wavelets decomposition

A. Identification of atrial fibrillation detection problems.

- No P waves, fibrillating chaotic F waves around baseline
- Approximate atrial rate: 350 – 600 [bpm] (or no regular pattern discernible),
- Ventricular rate: 100 – 180 [bpm]
- Ventricular rhythm: Irregular; R-R interval constantly varies, as does size of QRS complexes

A library of clinical expert evaluated typical atrial fibrillation evolutions (example presented in Fig.1) was created as a database for optimal matched wavelet construction.

Manuscript received June 24, 2011..

P. S. Kostka is with the Silesian University of Technology, Institute of Electronics, 16 Akademicka St. Gliwice, Poland (corresponding author phone: +48 32 2371160; fax: +48 32 2372225; e-mail: pkostka@ polsl.pl).

E. J. Tkacz is with the Silesian University of Technology, Institute of Electronics, 16 Akademicka St. Gliwice, Poland (e-mail: etkacz@polsl.pl).

Fig.1 12 lead atrial fibrillation events (indicative of atrial fibrillation with a fast ventricular response) as a source of reference pattern in created library for matched wavelet design.

B. Matched wavelet construction algorithm.

An algorithm introduced by Chapa et al. [5] was modified to obtain optimal wavelet basis function for AF detection. According to this optimized method, wavelet basis function matching is performed separately for magnitude and phase of the signal spectrum. To obtain a single wavelet that could provide the best match, the signal of interest is projected onto the different orthonormal wavelets.

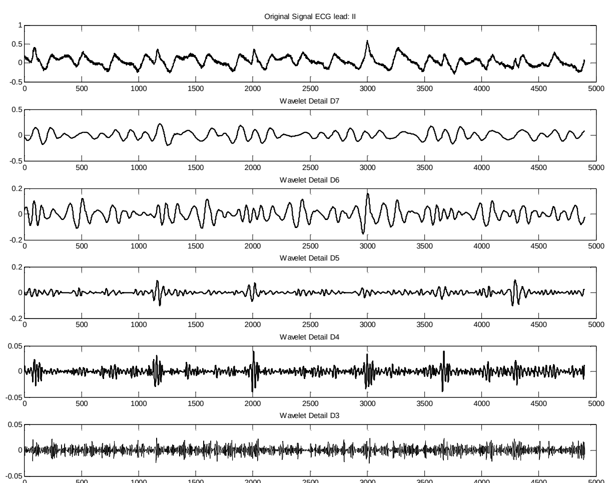

Fig.2 An example of multilevel Mallat decomposition components (details: d3-d7 with original ECG signal) of II ECG lead of patient with AF.

C. Classifier system structure with matched wavelet analysis used as a feature extraction tool.

Before a normal performance of presented classifier (Fig.3 'left channel') a matched wavelet is created for a library of representatives AF templates (Fig.3 'right channel'). The optimized basic wavelet is a crucial point of Time-Frequency analysis of ECG signal in order to create the most representative feature vector, consisting of set of energies $E\{S_{(j)}\}$ and entropies $E\{S_{(j)}\}$ of chosen ECG wavelet analysis decomposition components ($j=n..m$), which correspond to filter bank with definite frequency ranges. So instead of original ECG raw data a lower dimension feature vector *F2* is put as an input to final classifier.

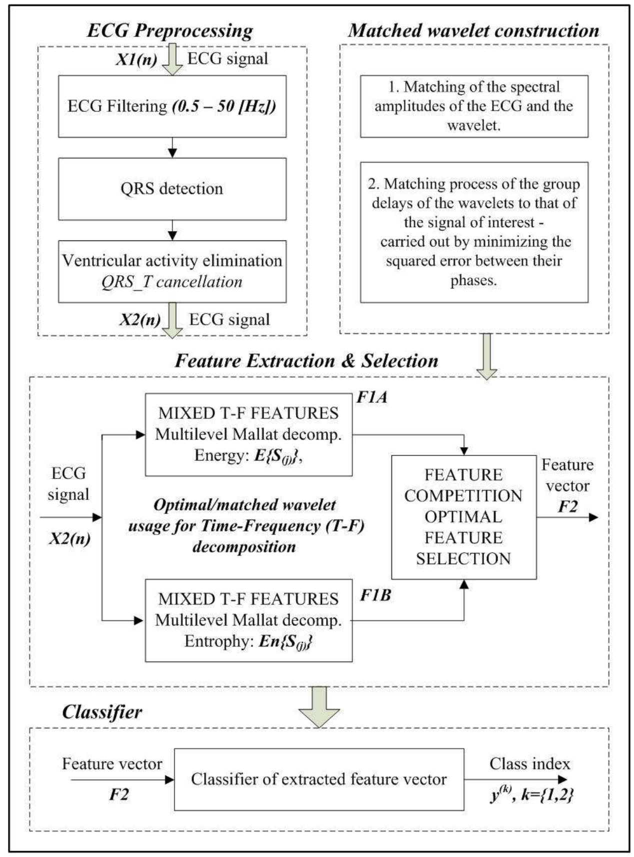

Fig.3 Structure of presented method with a matched wavelet construction (right upper 'channel') and normal AF detection from ECG data path (left upper 'channel').

D. ECG signal preprocessing with cancellation of ventricular activity.

On the first preprocessing stage of input ECG signal analysis apart from standard ECG filter implementation the elimination of ventricular activity by QRS complex and T-wave cancellation determining the quality of whole procedure was carried out [8][9]. Literature review of papers connected to AF detection problem shows positive influence of ventricular activation cancelation by removing QRST complex from original signal for further analysis [10]. Traditional methods based on simple QRS cycles averaging and subtraction was not enough in case of significant beat-to-beat changes in real ECG signal, which cannot be treated as periodical signal. Recent studies report principal component analysis, blind source separation and artificial neural networks [11] as a alternative promising tools for QRST complex cancellation. In presented approach we proposed to use the advantages of discrete wavelet transform analysis dedicated for non-stationary signals. Reconstructed with threshold detail components ***d4+d5*** of multilevel Mallat decomposition created the base for QRS detection and cancellation by subtraction from remain components.

E. Support Vector Machine as an extracted features classifier

1) SVM structure. The support vector machine (SVM) is a promising classification technique proposed by Vapnik and his group at AT&T Bell Laboratories [12]. SVM is a good tool for the two classifications. It can separate the classes with a particular hyperplane which maximizes a quantity called the margin. The margin is the distance from a hyperplane separating the classes to the nearest point in the dataset. The advantage of maximum margin criterion is its robust characteristic against noise in data, and making a solution unique for linearly separable problems. In addition, it is important that the SVM with a theoretically strong support is based on the statistical learning theory framework.

2) SVM kernel selection. Matching basic kernel function for specified task is remarkably for SVM classifier performance. Literature described investigations [13] indicate a polynomial kernel as a good solution for prediction problems [14] while radial basis function kernel better suited for classification [15]. Two mentioned types of SVM kernels, expressed by formulas (4) , (5) were used in presented paper:

a) Polynomial Kernel of *d* - degree

$$k_P = k(x_i, x') = \left(a + b\langle x_i \cdot x' \rangle\right)^d \qquad (4)$$

with a and b as a constants. A special case for a=0 ; b=d=1 forms a linear kernel.

b) Gaussian Radial Base Function (RBF) Kernel

$$k_G = k(x_i, x') = \exp\left(-\frac{1}{\gamma}\left\|x_i - x'\right\|^2\right) \qquad (5)$$

with γ width specified a priori.

3) Nonlinear case:

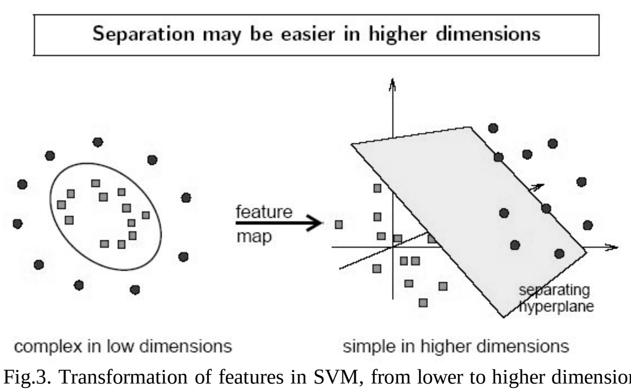

Fig.3. Transformation of features in SVM, from lower to higher dimension space, where the classification can be easier

SVM to effectively handle real-world data, often in non-linear dependences allow to model a non-linear decision surfaces. According to method proposed in [10], the idea is to explicitly map the input data to some higher dimensional space, where the data is linearly separable (fig.2). We can use a mapping (3):

$$\Phi : \Re^N \to \Im \qquad (3)$$

where *N* is the dimension of the input space, and $\Im$ a higher-dimensional space, termed *feature space*. In feature space, the technique described in the above section can be used to find an optimal separating hyperplane. When the hyperplane is found, it can be mapped back down to input space. If a non-linear mapping Φ is used, the resulting hyperplane in input-space will be non-linear.

III. RULES EXTRACTION ALGORITHM FOR SVM AND NEURAL NETWORK CLASSIFIER STRUCTURES

The TREPAN algorithm developed by Craven [13] is a relatively novel rules extraction method converting the knowledge gathered in black-box type structure like e.g. SVM or Neural Networks in IF-THEN rules. On the first stage TREPAN extracts a decision trees from parameters of black-box models using a concept of recursive partitioning similar to induction algorithms. TREPAN forms trees that use M-of-N expressions. An M-of-N expression is satisfied when at least m of its n conditions are satisfied.

IV. RESULTS

A. Feature extraction by means of wavelet decomposition

Selection of appropriate wavelet and the number of decomposition levels is very important in analysis of signals using the wavelet transform. Usually, tests are performed with different types of wavelets and the one which gives maximum efficiency is selected for the particular application. In this work we designed an own wavelet based on best matching algorithm for problem of atrial fibrillation (AF) detection. A created matched wavelet was compared with often used as a reference basic function the Daubechies wavelet of order 2, with the smoothing feature making it more suitable to detect changes of the ECG signals. Therefore, the final wavelet coefficients as a base for new feature vector creation were computed using the matched AF wavelet.

To verify presented method, ECG signals taken from MITBIH database containing AF episodes were tested. Whole data set consisting of 40 cases with long term ECG recordings were divided into learning and verifying set. Optimal number of selected features for classifier effectiveness is presented in fig.4. Performance of presented pattern recognition system was evaluated based on classical measures of classifier Sensitivity and Specificity.

B. SVM Model of black-box classifier

The Matlab implementation of Support Vector Machine theory was used to train the SVM and to adjust its parameters, which was on the next stage put to SVM-TREPAN in order to extract the rules gathered inSVM structure during the learning phase. To select the best set of parameters (γ,C) for the Gaussian kernel, the grid-search and cross-validation approach with $\gamma=[2^8,2^7,2^6,..,2^{-8}]$ and $C=[2^{10},2^9, 2^8,..,2^{-4}]$ was used. Different pairs (γ,C) were tested and the one with the best cross-validation accuracy was selected.

For evaluation of new feature set quality based on matched wavelet decomposition, the performance of different type of classifier structures with chosen type of SVM kernel function with reference systems with no feature extraction (FE) stage were compared (TABLE I).

TABLE I
RESULTS OF DIFFERENT FEATURE EXTRACTION APPROACH IN CLASSIFIER STRUCTURES

CLASSIFIER STRUCTURE TYPE	SENSITIVITY	SPECIFICITY
Classifier without FE (*Polyn. kernel*)	0.65	0.62
SVM Classifier + db4 (*Gaussian kernel*)	0.71	0.73
SVM Classifier + Bior 2.4 (*Polyn. kernel*)	0.80	0.82
SVM Classifier + Matched Wavelet (*Gaussian. kernel*)	0.87	0.86

The most important feature characterizing classifier performance is its generalization ability. The measure of Sensitivity (SN) and Specicity (SP) was calculated for the chosen in previous subsection structure of whole SVM classifier with feature extraction stage. Apart from classification performance Table I presents also the influence of proposed feature extraction stage for the values of SN and SP obtained for the SVM classifier with no feature extraction stage.

V. CONCLUSIONS

Selection of appropriate wavelet and the number of decomposition levels is very important in analysis of signals using the wavelet transform. Usually, tests are performed with different types of wavelets and the one which gives maximum efficiency is selected for the particular application. In this work we designed an own wavelet based on best matching algorithm for problem of atrial fibrillation (AF) detection. Results indicate that a matched wavelet, that was able to capture the broad ECG features, could be obtained. Such a wavelet could be used to extract ECG features such as QRS complexes and P&T waves with greater accuracy.

Many applications of signal representation, adaptive coding and pattern recognition require wavelets that are matched to a signal of interest. Presented algorithm can be considered as a universal method, which can be applied to different types of biomedical signals by adjusting the process of matched wavelet creation, conditioned by the specific character of studied signals.

After preliminary data processing including important stage of ventricular activity cancellation a feature extraction from T-F domains based on designed wavelet, matched for atrial fibrillation detection problem was carried out. A SVM based structure was used to classify the new feature characterizing the analysed problem of atrial fibrillation detection. Presented article focuses on improving the SVM classification abilities by preliminary selecting features with maximal weight in classification process. It allowed to find the optimal feature subset selection of from different domain T-F features. Atrial Fibrillation detector tests gave for the optimal structure the value of classifier sensitivity SN=87%, while specificity SP=86% for AF with different degree of organization (atrial flutter, AF1, AF2 and AF3).

To conclude, obtained results showed, that before pattern classifier can be properly designed and effectively used, it is necessary to consider the feature extraction and data reduction problems. Feature extraction should consists in choosing those features, which are most effective for preserving the class separability. Support Vector Machine structure appeared as an effective tool for biomedical data classifier, improving whole classification process.

REFERENCES

[1] I. Daubechies Ten Lectures on Wavelets. Philadelphia, PA: SIAM, 1992, ch. 3-5.

[2] S. Mallat A theory for multiresolution signal decomposition, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[3] S.G. Mallat, C. Zhang Matching pursuit with time-frequency dictionaries, IEEE Trans. Signal Processing, vol. 41, pp. 3397–3415, Dec. 1993.

[4] S. Chen, D. Donoho Basis pursuit. Proc. IEEE 28th Annu. Asilomar Conf. Signals, Syst., Comput., Nov. 1994.

[5] M. Chapa, R. Rao Algorithms for designing wavelets to match a Specified signal, IEEE Transactions on Signal Processing, 2000, pp. 3395-3406.

[6] M. Unser, A. Aldroubi A review of wavelets in biomedical applications *Proc. IEEE*, vol. 84, no. 4, pp. 626–638, Apr. 1996.

[7] A.H. Tewfik, D. Sinha, P. Jorgensen On the optimal choice of a wavelet for signal representation, *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 747–765, Mar. 1992.

[8] Sanchez C., Millet J., Rieta J. J., Rodenas J., Castells F., Ruiz R., and Ruiz V.,Packet Wavelet Decomposition: An Approach For Atrial Activity Extraction, IEEE Computers in Cardiology, vol.29 Sept.2002

[9] C. Vasquez, A. Hernandez, F. Mora, G. Carrault, G. Passariello, ‘Atrial activity enhancement by Wiener filtering using an artificial neural network’, *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 8, pp. 940-944, Aug.2001.

[10] C. Vasquez, A. Hernandez, F. Mora,G. Carrault, G. Passariello. Atrial activity enhancement by Wiener Filtering using an artificial neural network. IEEE Transactions on Biomedical Engineering, vol. 48, no. 8, pp. 940-944, Aug.2001

[11] P.S. Kostka, E.J. Tkacz. Feature extraction optimization in neural classifier of heart rate variability signals, Proceedings of the 4th International Conference on Computer Recognition Systems CORES 2005, Advances in Soft Computing, Springer-Verlag, pp.585-594, 2005.

[12] V. Vapnik. The Nature of Statistical Learning Theory. Springer, N.Y., 1995

[13] E. Osuna, R. Freund, and F. Girosi. Training support vector machines: An application to face detection. In Proceedings of CVPR’97, Puerto Rico, 1997

[14] B. Scholkopf, C. Burges, and V. Vapnik. Extracting support data for a given task. In U. M. Fayyad and R. Uthurusamy, editors, Proceedings, First International Conference on Knowledge Discovery & Data Mining, Menlo Park, 1995.