Feature Extraction From Parametric Time–Frequency Representations for Heart Murmur Detection

L. D. Avendaño-Valencia,[1] J. I. Godino-Llorente,[2] M. Blanco-Velasco,[3,4] and G. Castellanos-Dominguez[1]

[1]Departamento de Ingeniería Eléctrica, Electrónica y Computación, Universidad Nacional de Colombia, Km. 9, Vía al aeropuerto, Campus la Nubia, Caldas, Manizales, Colombia; [2]Departamento de Ingeniería de Circuitos y Sistemas, Universidad Politécnica de Madrid, Ctra. de Valencia, km. 7, 28031 Madrid, Spain; [3]Universidad de Alcala, Campus Universitario, 28871 Alcala de Henares, Madrid, Spain; and [4]Departamento Teoría de la Señal y Comunicaciones, Ctra. Madrid-Barcelona, km 33.628805, Alcalá de Henares, Madrid, Spain

Abstract—The detection of murmurs from phonocardiographic recordings is an interesting problem that has been addressed before using a wide variety of techniques. In this context, this article explores the capabilities of an enhanced time–frequency representation (TFR) based on a time-varying autoregressive model. The parametric technique is used to compute the TFR of the signal, which serves as a complete characterization of the process. Parametric TFRs contain a large quantity of data, including redundant and irrelevant information. In order to extract the most relevant features from TFRs, two specific approaches for dimensionality reduction are presented: feature extraction by linear decomposition, and tiling partition of the *t–f* plane. In the first approach, the feature extraction was carried out by means of eigenplane-based PCA and PLS techniques. Likewise, a regular partition and a refined Quadtree partition of the *t–f* plane were tested for the tiled-TFR approach. As a result, the feature extraction methodology presented, which searches for the most relevant information immersed on the TFR, has demonstrated to be very effective. The features extracted were used to feed a simple *k-nn* classifier. The experiments were carried out using 45 phonocardiographic recordings (26 normal and 19 records with murmurs), segmented to extract 548 representative individual beats. The results using these methods point out that better accuracy and flexibility can be accomplished to represent non-stationary PCG signals, showing evidences of improvement with respect to other approaches found in the literature. The best accuracy obtained was 99.06 ± 0.06%, evidencing high performance and stability. Because of its effectiveness and simplicity of implementation, the proposed methodology can be used as a simple diagnostic tool for primary health-care purposes.

ABBREVIATIONS

2D-PCA	Two-dimensional PCA
AR	Autoregressive
BIC	Bayesian information criterion
CWT	Continuous wavelet transform
ECG	Electrocardiogram
HS	Heart sound
k-nn	*k*-nearest neighbors
LS-TVAR	Least-squares TVAR
PCA	Principal component analysis
PCG	Phonocardiogram
PLS	Partial least squares
SNR	Signal-to-noise ratio
t–f	Time–frequency
TFR	Time–frequency representation
TVAR	Time-varying autoregressive
WVD	Wigner–Ville distribution

INTRODUCTION

Cardiac mechanical activity is appraised by auscultation and processing of heart sound (HS) records (known as phonocardiographic signals—PCG), which is an inexpensive and non-invasive procedure. Since computer-based analysis of HS may contribute to improve diagnosis of cardiac malfunctions, PCG has preserved its importance in many medical fields of

Address correspondence to L. D. Avendaño-Valencia, Departamento de Ingeniería Eléctrica, Electrónica y Computación, Universidad Nacional de Colombia, Km. 9, Vía al aeropuerto, Campus la Nubia, Caldas, Manizales, Colombia. Electronic mail: ldavendanov@unal.edu.co, cgcastellanosd@unal.edu.co, ldavendanov@gmail.com, igodino@ics.upm.es, manuel.blanco@uah.es

clinical practice.[2,10,19] Specifically, murmurs, which are sorted into systolic or diastolic, are some of the basic signs of pathological changes to be identified, but they overlap with the cardiac beat, and these HS cannot be easily separated by the human ear, since a large amount of information in the frequency domain is below the audibility thresholds. Moreover, cardiac murmurs are non-stationary signals that exhibit sudden frequency changes and transients. Therefore, the time–frequency representations (TFRs) have been proposed before to investigate the correlation between the time–frequency (t–f) characteristics of murmurs and the underlying cardiac pathologies.[19] For that matter, in the literature, both non-parametric and parametric estimations of TFR are generally employed.[7,15,17] Non-parametric methods (e.g., Wigner–Ville distribution—WVD, continuous wavelet transform—CWT, etc.) are based on non-parameterized representations of the signals as a simultaneous function of time and frequency, whereas parametric methods are based on parameterized expressions of a time-dependent autoregressive modeling (or related types) and their extensions. In biomedical applications, the non-parametric methods of estimation that are commonly implemented by means of wavelet transform suffer of a trade-off between time and frequency resolution.[21]

Since there are large differences in the transition patterns among the individual sets of multi-frame signals, a stable estimation of the transition patterns should be carried out. In this line, and due to its intrinsic generality and its capacity to detect formant frequencies, the time-varying autoregressive (TVAR) models, in which the auto-regressive coefficients are allowed to vary with time, have provided useful empirical representations of non-stationary time series for biomedical signal analysis.[6] The frequency resolution of the parametric methods is superior because of the implicit extrapolation of the autocorrelation sequence.[21] Furthermore, the TVAR technique can provide a higher resolution with respect to the non-parametric estimations, and without the complication of the quadratic terms regarded to the quadratic TFR, or the need to generate a high time-resolution wavelet analysis. Despite these advantages, the selection of the model order and the methods of estimation for the parameters are the two main problems, which had discouraged the use of TVAR models to estimate the TFR.[6] Explicitly, the selection of the model order has a large effect on the quality of the signal representation,[15] and for pathological PCG signals it is often difficult to select a unique value correctly.[12,14,16] As a result, the TVAR model may not resolve well enough the fine structure in the data. However, the methods of estimation of the TVAR parameters can help with this

issue. In general, these methods should be directly coupled with the time-varying dynamics of the PCG signal. Assuming that the signal is stationary within short segments, and under the assumption that the parameters and innovations variance are independent of time,[17] AR estimators can be used in short time windows. Because of the lack of resolution, this method is basically suitable for cases where the evolution in the dynamics is slow. This is the case of the PCG signals. It must be quoted that regardless of the improvement of the TVAR models for the representation of non-stationary signals, the modeling of the PCG signal to automatically detect murmurs remains still an open issue.

For the automatic detection of murmurs, the TFR-based classification methods are preferred with respect to other techniques because t–f planes have clear discriminant capabilities to separate signals belonging to different classes. In addition, even that the flexibility to form the feature vector in 2D representations is considered the main advantage of a t–f domain-based classification, this enhanced representation also tends to contain a large amount of redundant data, and hence a dimensionality reduction is required. Thus, there is a growing need for new data reduction methods that can accurately parameterize the activity in TFR of biosignals, particularly those with higher resolution, as stated in Bernat et al.[5] for the case of electroencephalographic signals. A direct approach is the use of principal component analysis (PCA) to reduce the dimensionality of the feature space resulting from the t–f plane, assuming that the information is equally distributed along with the TFR. As concluded in Englehart et al.,[11] a PCA-reduced t–f representation has shown to effectively accommodate the loosely structured waveforms of some transient biological signals, when quantitative and decision-based analysis is required. PCA models each TFR as the weighted sum of base functions obtained as the components, which maximize the variability on the dataset. Nevertheless, for classification purposes, the components obtained are not always related with the most discriminative information. Thus, similar decomposition approaches, such as partial least squares (PLS), can be used as a supervised dimensionality reduction approach that yields components maximally related with labels.[4]

Nonetheless, as commonly known in the state-of-the-art, the methods, which classify the TFRs based on the local regions of the t–f plane, have achieved higher success rates than those based on the entire image[22]; and for PCG signals this supposition is more adequate, since most of the information is concentrated wherever the heart murmurs are present. But there is a noteworthy unsolved issue associated with local-based analysis, namely, the selection of the size of local

relevant regions. As a result, the choice of feature extractors in the *t*–*f* domain, and the classifier, is highly dependent on the final application.[19]

A major motivation in this study is to generate a set of parametric TFR-based features extracted from PCG recordings, capable of detecting murmurs with higher accuracy than using static features. So, the aims of this study are (a) the calculation of TFRs using a TVAR model, taking directly into account the variability of the PCG induced by the murmurs originated by valve pathologies; (b) the evaluation of the best set of dynamic features, estimated from parametric-based TFR, and suitable for the classification of heart murmurs.

Throughout this article, the criterion followed to compare the different approaches is the classifier accuracy, obtained using the well-known *k-nn* approach. For the sake of comparison and for the different feature extraction techniques described, the results—in terms of accuracy—of the proposed parametric approach were compared with those obtained using two non-parametric methods, namely, WVD and CWT (estimated as explained in Quiceno-Manrique *et al.*[18]). Moreover, the results were compared with other assumed baseline methods in the literature.[8,23]

The rest of this article is organized as follows: first, the time-varying representation implemented is introduced, followed by a brief description of the TVAR parameter estimation method used, computed by minimizing the mean-square error. Then, the methodology for feature extraction is described in detail (based on linear decomposition and partition schemes). Lastly, the effectiveness of a feature set based on parametric TFR representing the dynamics of the HS is applied to the detection of murmurs, comparing the results with other methods in the literature. Finally, there is a brief discussion of the results obtained.

BACKGROUND

The stochastic parameters that define the TVAR model, and that are used to generate the TFRs, are briefly described in this section. Moreover, the feature extraction techniques from TFRs considered in this article are presented.

TVAR Models with Locally Stationary Assumptions

A time-variant autoregressive model of *p*th order, shortly TVAR(*p*), is described as follows

$$x[t] = \boldsymbol{\theta}^{\mathrm{T}}[t]\mathbf{h}[t-1] + e[t]; \quad e[t] \sim N(0, \sigma_e^2[t]) \quad (1)$$

where $\mathbf{h}[t] = \{x[t-i] : i = 1, \ldots, p\}$, $\mathbf{h} \subset \mathbb{R}^p$ is a non-stationary (real-valued vector) processes to be

modeled, $e[t]$ is an unobservable and uncorrelated Gaussian sequence with zero mean and time-dependent variance $\sigma_e^2[t]$ (or innovations variance) and $\boldsymbol{\theta}[t] = \{\theta_i[t]\}$, $\boldsymbol{\theta} \subset \mathbb{R}^p$ are the parameters of the TVAR model. The order of a stationary model can be used as starting point to set the TVAR model order,[1] but a more accurate estimation can be achieved by trial-and-error procedures using a fitting function such as Bayesian information criterion (BIC).[17]

The vector of parameters, $\boldsymbol{\theta}[t]$, and the innovations variance, $\sigma_e^2[t]$, for a given value of the sampling frequency, f_s is related to the spectral content of $x[t]$ by[21]:

$$\mathbf{S}_x(t,f) = \frac{\sigma_e^2[t]}{\left|1 + \sum_{i=1}^{p} \theta_i[t]e^{-j\omega it/f_s}\right|^2}, \quad \mathbf{S}_x(t,f) \subset \mathbb{R}^{T \times F} \quad (2)$$

that can be assumed as the time-varying power spectral density (i.e., the TFR) of the signal if the system were stationary at the time instant *t*. The dimension of the TFR is given by the time and frequency resolution (*T* for time, and *F* for frequency).

TVAR representations differ from their conventional stationary counterparts in that they are time-dependent. The methods upon them provide a set of potential advantages listed in Poulimenos and Fassois,[17] namely, (a) its simplicity of representation, since these models may be potentially specified by a limited number of parameters; (b) improved accuracy; (c) better resolution; (d) better tracking of the time-varying dynamics; (e) flexibility of analysis, since the parametric methods are capable of directly capturing the underlying structural dynamics regarded to the non-stationary behavior.

The TVAR model of Eq. (1) is completed specifying the evolution of $\{\theta_i[t]\}$ and $\sigma_e^2[t]$. These two parameters will take a value at the time instant *t* for each window of analysis. Nonetheless, if the coefficients at each time are assumed to be independent, then the number of parameters needed becomes usually greater that the amount of data. In the context of biological signal processing and to overcome this problem, the stochastic regression approach is likely more suitable.[6] Specifically, in a locally stationary method for TVAR model estimation (referred as LS-TVAR), the parameters, $\boldsymbol{\theta}[t]$, are commonly computed by minimizing the mean-square error within a window of size *M*:

$$\widehat{\boldsymbol{\theta}}[t] = \arg\min_{\boldsymbol{\theta}} \sum_{\tau=t-M}^{t} \left(x[t] - \boldsymbol{\theta}^{\mathrm{T}}[t]\mathbf{h}[t-1]\right)^2 \quad (3)$$

leading to the Yule-Walker equations that can be solved by any of the standard algorithms such as

Burg's one.[21] The tracking ability of the LS-TVAR algorithm is prescribed by either the size of a window or the value of a fading rate, but conditioning the limitation of this approach: if the process evolves too quickly, the algorithm will not properly track the evolution of the coefficients, unless the window becomes very short which in turn degrades the spectral resolution.[7] Consequently, for the LS-TVAR approach (Eq. 3), only a coarse adjustment of the tracking properties can be carried out,[13] which is enough for the slowly changing dynamics of the PCG signals.

In order to complete the LS-TVAR model, the variance of the observation noise $\sigma_e^2[t]$ must be estimated. A sliding window on the square of estimation error[21] can be used:

$$\sigma_e^2[t] = \frac{1}{M} \sum_{\tau=t-M}^{t} w[\tau, \alpha] e^2[t] \tag{4}$$

where $w[\tau, \alpha]$ is a smoothing Gaussian window with aperture of value α and length M which determines the quantity of data considered in the window.

Feature Extraction From the Enhanced t–f Planes

The TFR can be represented as a matrix, $\mathbf{S}_x(t,f) \subset \mathbb{R}^{T \times F}$, with T time and F frequency points. This matrix can be used to characterize the dynamics of the PCG signals, but contain a large amount of information difficulting further processing. Thus, the problem stands on how to reduce the large amount of information and redundancies of the TFRs, but keeping most of the relevant information. The goal is to reduce the original TFR, $\mathbf{S}_x(t, f)$, into a set of features, $\xi \subset \mathbb{R}^n$, with a smaller dimension but containing the most important information available in the former dataset. The main drawback is that TFRs lie on a bidimensional space, $\mathbb{R}^{T \times F}$, while a set of features amenable for a simple classifier should reside on an unidimensional space $\mathbb{R}^n$.

A simple approach for manipulating the TFR matrix is reducing its dimension by downsampling or even decimating the time and frequency resolution. However, this method has not been considered, since it removes some important information from the TFR. An alternative approach to reduce the dimensionality would be to divide the two-dimensional TFR in two-dimensional bins with similar characteristics, characterizing each bin with an average of the energy contained within it. The bigger the area of these bins, the smaller the final feature space, but the lower the resolution. These bins could be equally distributed or defined according to the amount of information present in each part of the TFR. On the other hand, as quoted before, another direct approach for feature extraction would be the use of a linear decomposition method (e.g., PCA, 2D-PCA, PLS) to decrease the dimensionality projecting the whole TFR into a smaller subspace but keeping the most relevant information. The feature extraction approaches used throughout this study have been grouped into two subsections: (a) tiled-TFR approaches; and (b) linear decomposition-based approaches.

Tiled-TFR Approaches

A straightforward approach for 2D feature extraction is to divide a given t–f plane into tiles evaluating their informativity (the lower the probability of the dynamic occurrence, the higher the tile informativity). The informativity can be carried out computing a statistic operator, $\mathcal{L}\{\cdot\}$, over a determined tile of the t–f plane $\mathbf{S}_x(\Delta t, \Delta f)$ such as the variance $\text{var}\{\cdot\}$:

$$\xi = \text{var}\{\mathbf{S}_x(\Delta t_i, \Delta f_j)\}; \quad i = 1, \ldots, N_{\Delta t}; \; j = 1, \ldots, N_{\Delta f} \tag{5}$$

where Δt_i stands for the ith time segment; Δf_j is for the jth frequency bin; $N_{\Delta f}$ stands for the number of frequency bins; and $N_{\Delta t}$ is the number of time segments. Therefore, it is expected that the feature set extracted from Eq. (5), ξ, holds enough information related to the non-stationary properties of the signal.

In the beginning, the t–f tiles can be made up of fixed size frameworks, so that the feature vector can be described as $\xi = \{\xi_i : i = 1, \ldots, N_b; \; \xi_i \in \mathbb{R}^1\}$, $\xi \in \mathbb{R}^{N_b}$, where $N_b = N_{\Delta f} \cdot N_{\Delta t}$ is the number of tiles. Both, $N_{\Delta f}$ and $N_{\Delta t}$ can be determined empirically.[9,22] Time and frequency partitions are defined as:

$$\Delta t_i = [\, t_{i-1} \quad t_i \,], \quad \text{where } t_i = \frac{i \cdot T}{N_{\Delta t}}, \quad \forall i = 1, \ldots, N_{\Delta t}$$

$$\Delta f_j = [\, f_{j-1} \quad f_j \,], \quad \text{where } f_j = \frac{j \cdot F}{N_{\Delta f}}, \quad \forall j = 1, \ldots, N_{\Delta f} \tag{6}$$

This regular configuration of time windows and frequency bins assumes that the information content is equally distributed along the t–f plane. However, for PCG signals, this assumption might be inadequate, since most of the information is concentrated around the HS,[1] S1, and S2, and eventually between them, in the systole and diastole, wherever the heart murmurs are present.

---

[1]S1 implies the closing of the tricuspid and mitral valves immediately preceding the systole, whereas S2 corresponds to the closing of the aortic and pulmonary valves at the end of systole.

Thus, a more efficient representation might be reached using unfixed and adaptive size frameworks: for instance, using adaptive multiscale representations via the *Quadtree decomposition* that splits the t–f domain into four equally sized tiles, where each tile can be successively decomposed into four new tiles.[20] The root of the tree is the initial TFR.

Linear Decomposition-Based Approaches

The PCA and the PLS methods were used throughout this article as unsupervised and supervised methods, respectively, to perform dimensionality reduction in the TFRs. Moreover, a two-dimensional extension of PCA was used to represent the reduced feature space.

PCA Decomposition

Let $\mathbf{X} = \{\boldsymbol{\chi}_i: \; i = 1,\dots,N_r\}$ be a set of N_r objects generated by m random variables. Thus, for the ith object, the respective data set $\boldsymbol{\chi}_i = \left[\chi_i^1, \chi_i^2, \dots, \chi_i^m\right]$ is given, and from which the centralized data matrix is built as:

$$\mathbf{X}_0 = \left[(\boldsymbol{\chi}_1 - \boldsymbol{\mu}_\chi)^{\mathrm{T}}, (\boldsymbol{\chi}_2 - \boldsymbol{\mu}_\chi)^{\mathrm{T}}, \dots, (\boldsymbol{\chi}_{N_r} - \boldsymbol{\mu}_\chi)^{\mathrm{T}}\right]^{\mathrm{T}},$$

$$\boldsymbol{\mu}_\chi = \frac{1}{N_r}\sum_{i \in N_r} \boldsymbol{\chi}_j, \quad \mathbf{X}_0 \in \mathbb{R}^{N_r \times m} \tag{7}$$

The conventional PCA looks for an orthogonal transformation matrix $\mathbf{W}$, (being $\mathbf{W}^{\mathrm{T}}\mathbf{W} = \mathbf{I}$, $\mathbf{I} \in \mathbb{R}^{n \times n}$, $\mathbf{W} \in \mathbb{R}^{m \times n}$), to project the data onto a smaller set of variables with the maximum variance, by means of the linear transformation $\mathbf{Y} = \mathbf{X}_0\mathbf{W}$, where:

$$\mathbf{W} = \arg\max_{\mathbf{W}} \mathrm{tr}(\mathbf{W}^{\mathrm{T}}\mathbf{X}_0^{\mathrm{T}}\mathbf{X}_0\mathbf{W}) \tag{8}$$

In practice, the solution is found by setting the columns of $\mathbf{W}$ to the n leading eigenvectors of the covariance matrix $\mathbf{X}_0^{\mathrm{T}}\mathbf{X}_0$.

Now consider the case when each object is not defined as a vector but as a matrix, such as TFR matrices. Then, each object is described by a dataset $\mathbf{S}_{x_i}(t,f) \subset \mathbb{R}^{T \times F}$, which from now on will be denoted by $\mathbf{S}_i$ for shorter notation. Every variable of each set $\mathbf{S}_i$ is time-dependent and has been measured upon a set of T instants of time. Thus, for each object the following dataset is given $\left\{\mathbf{S}_i^{(j,k)} : i = 1,\dots, N_r; j = 1,\dots, T; k = 1,\dots, F\right\}$, where notation $\mathbf{S}_i^{(j,k)}$ stands for the kth point, measured for the ith object, at the instant of time j.

PCA can be carried out on these data performing the same singular value decomposition but on a data matrix of vectorized t–f domains, thus dealing with the stochastic nature of the variables by assuming that

each instant of time and frequency point $S_i^{(j,k)}$, $\forall j$, constitutes a new random variable. Therefore, each object can be described by

$$\boldsymbol{\chi}_i = \left[S_i^{(1,1)}, \dots, S_i^{(1,F)}, \cdots, S_i^{(T,1)}, \dots, S_i^{(T,F)}\right],$$

$$\boldsymbol{\chi}_i \in \mathbb{R}^{T \times F} \tag{9}$$

and a linear component decomposition can be carried out over the rewritten centralized data matrix in Eq. (7). Lastly, it must be quoted that the PCA transformation provides a means of unsupervised dimensionality reduction, as no class membership qualifies the data when specifying the eigenvectors of maximum variance.

2D-PCA Decomposition

Another approach of bi-dimensional component analysis (known as 2D-PCA) is discussed in Yang *et al.*[24] and Avendano-Valencia *et al.*[3] Further refinement of the object description represented in Eq. (9) can be achieved if the vector used to represent an object, $\boldsymbol{\chi}_j \in \mathbf{X}$, is represented by the following matrix, taking into account the variability of the whole variable set:

$$\boldsymbol{\chi}_i = \begin{bmatrix} S_i^{(1,1)} & S_i^{(1,2)} & \cdots & S_i^{(1,F)} \\ S_i^{(2,1)} & S_i^{(2,2)} & \cdots & S_i^{(2,F)} \\ \vdots & \vdots & \ddots & \vdots \\ S_i^{(T,1)} & S_i^{(T,2)} & \cdots & S_i^{(T,F)} \end{bmatrix}, \quad \boldsymbol{\chi}_i \in \mathbb{R}^{T \times F} \tag{10}$$

In this case, the matrix of the data projected, $\mathbf{Y} = [\boldsymbol{\psi}_1^{\mathrm{T}}, \dots, \boldsymbol{\psi}_{N_r}^{\mathrm{T}}]^{\mathrm{T}}$, is described by the elementary matrices $\boldsymbol{\psi}_i = \boldsymbol{\chi}_i\mathbf{V}$, $\boldsymbol{\psi}_i \in \mathbb{R}^{T \times n_F}$. The reduction of the model in Eq. (8) is carried out over the column of the objects, which implies that the variables projected are capturing the variability of each object in time.

Also, the description of the object in Eq. (11) can be transposed to compute a transformation matrix $\mathbf{W} \in \mathbb{R}^{F \times n_T}$. Such matrix can reduce the dimensionality of the $\boldsymbol{\chi}_i$ rows, by means of the operation $\boldsymbol{\psi}_i = \mathbf{W}^{\mathrm{T}}\boldsymbol{\chi}_i$, where the matrix $\mathbf{W}$ is calculated from the matrix set $\{\boldsymbol{\varphi}_i = \boldsymbol{\chi}_i^{\mathrm{T}} : i = 1,\dots, N_r\}$. After the calculation of the arrangements, $\mathbf{V}_{T \times n_f}$, $\mathbf{W}_{F \times n_T}$, a column–row-based dimensionality reduction is carried out for each $\boldsymbol{\chi}_j$, i.e., $\boldsymbol{\psi}_i = \mathbf{W}^{\mathrm{T}}\boldsymbol{\chi}_i\mathbf{V}$, $\boldsymbol{\psi}_i \in \mathbb{R}^{n_T \times n_F}$. As a result, the dimensionality reduction takes into account not only the instant-by-instant variability of each random variable, given by the model represented in Eq. (9), but also checks for information variability through the frequency spectrum.

Mainly, the difference between PCA and 2D-PCA approaches is based on the form to tackle the matricial data. Both algorithms are graphically depicted in Fig. 1.
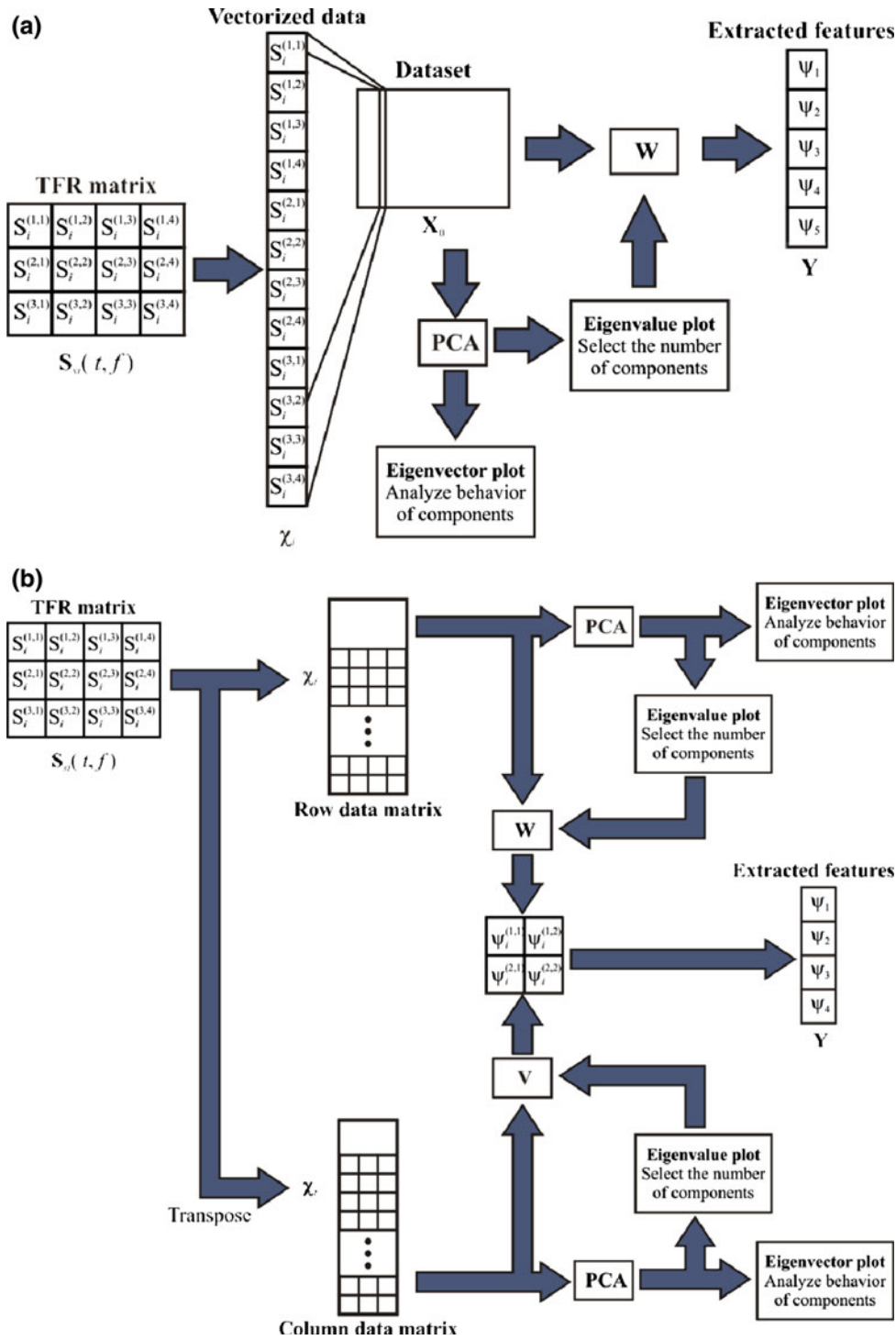
FIGURE 1. Dimensionality reduction procedure for TFR matrices; (a) PCA and vectorization; (b) 2D-PCA approach.

PLS Decomposition
This regression is a recent technique that somehow generalizes features from PCA, by projecting a set of dependent variables from a set of independent variables, but preserving the asymmetry of the relationship between independent variables, represented by a (very) large set, and dependent variables, whereas PCA treats them symmetrically. Specifically,

a multidimensional input TFR space $\mathbf{X}_0$, can be projected onto the basis vectors (planes), $\{\lambda_i: i = 1,\ldots,m\}$, through the following transformation:

$$\mathbf{w}_i = \lambda_i^T \mathbf{X}_0, \quad i = 1,\ldots,m \quad (11)$$

The set of weights $\boldsymbol{\Phi} = \{\mathbf{w}_i\}$ describes the contribution of each basis plane to the input TFR, and Eq. (11) is used as a feature vector for recognition tasks. The vector $\lambda_i = \text{diag}\{\Sigma\}$ comprises the diagonal elements of the diagonal matrix Σ, resulting after singular value decomposition of $\mathbf{X}_0$ (i.e., $\mathbf{X}_0 = \mathbf{U}\Sigma\mathbf{V}$, being $\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}$ the matrices of the left and right singular vectors). The PCA orthogonal decomposition of Eq. (11) remains the same for PLS, but choosing in a different way the latent vectors where additional conditions are required. So, the new basis vectors are recalculated after simultaneous decomposition: $\mathbf{X}_0 = \mathbf{T}\mathbf{P}^T + \varepsilon_X$, with $\mathbf{T}^T\mathbf{T} = \mathbf{I}$, and $\mathbf{Y} = \mathbf{T}\mathbf{Q}^T + \varepsilon_Y$, being the matrices ε_X and ε_Y the error terms, assumed to be i.i.d. normal. $\mathbf{P}$ and $\mathbf{Q}$ are the weight matrices used to reveal the influence of the individual variables $\mathbf{X}_0$ and $\mathbf{Y}$, respectively.

EXPERIMENTAL SETUP

Figure 2 shows the proposed methodology for heart murmur detection appraising the next stages: (a) preprocessing; (b) estimation of the parameters of the TVAR model; (c) computing the TFRs from the TVAR parameters; (d) feature extraction and dimensionality reduction of the *t–f* planes; and (e) training and validation.

Database Acquisition and Pre-Processing

The database used in this study is made up of 45 adult subjects, who gave their informed consent approved by an ethical committee, and underwent a medical examination. A diagnosis was carried out for each patient and the severity of the valve lesion was evaluated by cardiologists according to clinical routine. A set of 26 patients were labeled as normal, while 19 were tagged as pathological with evidence of systolic and diastolic murmurs caused by valve disorders. Furthermore, eight phonocardiographic recordings corresponding to the four traditional focuses of auscultation (mitral, tricuspid, aortic, and pulmonary areas) were taken per patient in the phase of postexpiratory and postinspiratory apnea. Every recording lasted 12 s approximately, and was obtained from the patient standing in dorsal decubitus position. Next, after visual and audio inspection by cardiologists, some of the eight signals were removed because of artifacts and undesired noise; moreover, it was taken into consideration that most of the time murmurs do not necessary show up for all focuses at once, unless they are very intense (which is evidence of their harmfulness). An electronic stethoscope (WelchAllyn® Meditron model) was used to acquire the PCG simultaneously with a standard 3-lead ECG (since the QRS complex is clearly determined, DII derivation is synchronized as a time reference). Both signals were sampled with 44.1 kHz rate and amplitude resolution of 16 bits. Tailored software was developed for recording, monitoring, and editing the HS and ECG signals.

Preprocessing was carried out as in Quiceno-Manrique *et al.*,[18] and consists on downsampling at 4000 Hz, amplitude normalization and inter-beat segmentation. Finally, the database holds 548 HS beats in total: 274 with murmurs (73 of diastolic class and 201 systolic), and 274 that were labeled as normal. The selection of the 548 beats used for training and validation was carried out by expert cardiologists, who choose the most representative beats of normal and pathologic patients (with murmurs). Table 1 summarizes the records stored in the database.

Although the database discriminates among normal, systolic, and diastolic murmurs, the objective in this study consists on the detection of heart murmurs for screening purposes.

Methods to Adjust the Parameters of the TFR Enhancement Techniques

Estimation of the TVAR Model Order

As stated before, an improved TFR estimation of non-stationary signals can be achieved by a locally

TABLE 1. Recordings stored in the database.

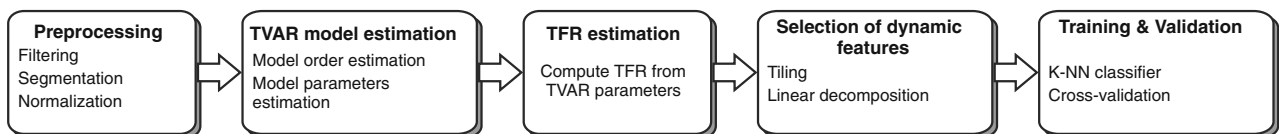|	Systolic	Diastolic	Normal	Total
Patients	8	11	26	45
Beats	201	73	274	548

FIGURE 2. Experimental outline of murmur detection, and the methods subject to investigation.

stationary TVAR model (LS-TVAR), defined by the model order p. The literature reports that a large order of the linear predictor (e.g., 28) is necessary to model the PCG signal over its full frequency range.[16] Nevertheless, the use of parametric TVAR modeling allows representing the process with a lower quantity of components.[17] But, as explained in Poulimenos A, Fassois,[17] a more accurate selection of the model order can be achieved based on the minimization of a *fitness* function. Specifically, the fitness function is estimated for each HS (i.e., heart beat) by means of the BIC. As suggested in Poulimenos A, Fassois,[17] the approach consists on evaluating BIC for every family of parameters, $e[t]$ and $\sigma_e[t]$, obtained for a given order of the model p:

$$\text{BIC}(p) = -\sum_{t=1}^{N}\left(\ln\sigma_e^2[t] + \frac{e^2[t]}{\sigma_e^2[t]}\right) \qquad (12)$$

where N is the number of signal samples. The BIC curves were computed for different cut-off frequencies (raw signal, 500 and 1000 Hz), and ranging the model order from 1 to 15 (Fig. 3a). Likewise, the respective histograms are depicted in Fig. 3b for all the records in the database. The optimum model order is the one that minimizes the mean BIC function (i.e., the elbow of the curve). The figure shows that the lower the cut-off frequency signal the smaller the order selected for the model. Initially, the model order selected is $p = 6$, in line with the order reported in Güler *et al.*[12] and Kanai *et al.*[14] for time-varying AR modeling of HS.

Later, in the "Results" section, a finer empirical adjustment is presented ranging p from 1 to 15 and trying to maximize the classifier accuracy.

Estimation of the Parameters of the TVAR Model

The parameters of the autoregressive model with smoothness priors were estimated by means of the aforementioned LS-TVAR approach (Eq. 3). Under the assumption that the AR parameters do not change quickly, the PCG recordings were divided initially into windows of short duration (100 samples). The coefficients of the LS-TVAR model were found using the Burg Algorithm.

Once the parameters of the TVAR model were estimated, the TFRs were computed by means of Eq. (2). Figure 4 illustrates an enhanced TFR for normal and pathologic HS, estimated using non-parametric methods (WVD and CWT which were estimated as explained in Quiceno-Manrique *et al.*[18]), and the LS-TVAR parametric method used in this study. The TFRs shown in Fig. 4 are the matrices of dimension $T \times F$, where F is the number of spectral components of the PCG signal, $f = [0, 400]$ Hz; and T is the number of discrete-time samples of each recording. This arrangement is intended to cover the full-time range and a broad range of frequencies.

Methods to Adjust the Feature Extraction Algorithms

The next three sections describe the criteria and methods followed to adjust the aforementioned feature extraction and dimensionality reduction algorithms: (a) First of all, it is described the procedure to remove the area of the spectral surface that does not contain relevant information; (b) Secondly, and regarding the feature extraction for the tiled-based approach, it is defined empirically the size and distribution of the t–f
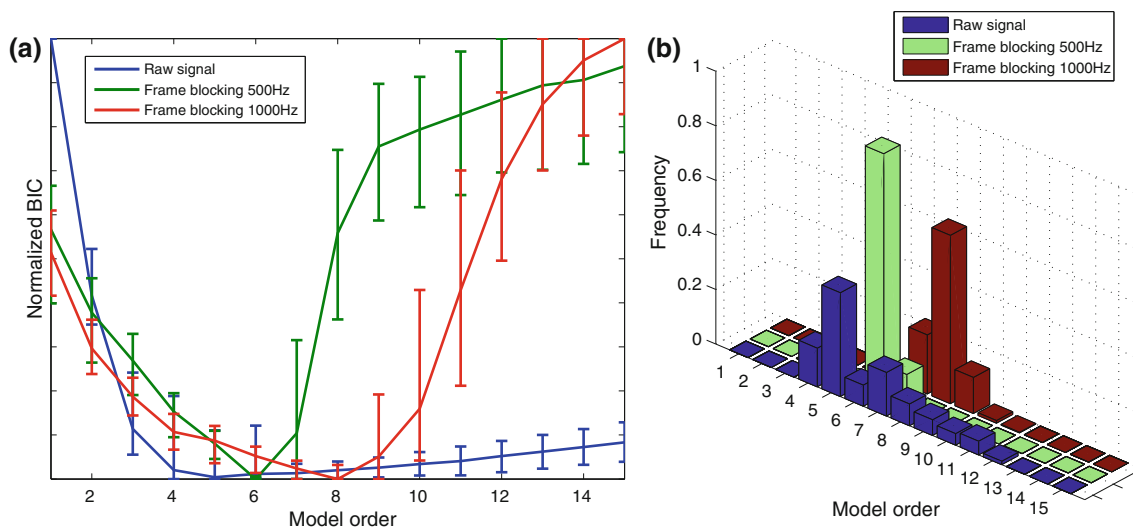

FIGURE 3. Estimation of model order for parametric TVAR model; (a) mean BIC curves; (b) histogram of minimum BIC values.

bins that divide the TFR; (c) Third, the methods to adjust empirically the size of the projected space with the linear decomposition methods (e.g., PCA, 2D-PCA, PLS) are described. The criterion used to adjust the feature extraction algorithms was the maximization of the classifier performance. A simple *k-nn* classifier was used in order to assess the goodness of the feature extraction methods.

Estimation of the Relevant Area of the TFRs

A straightforward observation of the TFR in Fig. 5 suggests that there are large areas of the spectral surface that do not contain relevant information for the detection of murmurs. Thus, the irregular concentration of information suggests cropping nil content areas that lie adjacent to the border of the *t–f* plane. Thus, a coarse estimation of the relevant area can be accomplished computing the complement set of any measure of informativity, given in Eq. (5):

$$\arg\left\{ \max \dim \left\{ \mathrm{var}\left(\mathbf{S}_x^{(0)}\right) \cup \mathrm{var}\left(\mathbf{S}_x^{(1)}\right) \right\} \right\} \geq \delta \quad (13)$$

where $\mathrm{var}(\mathbf{S}_x^{(k)})$, $k \in \{0, 1\}$ stands for the variance matrix of both TFR classes: normal and pathologic, respectively. The threshold, $\delta_{\min}$, is assumed to be
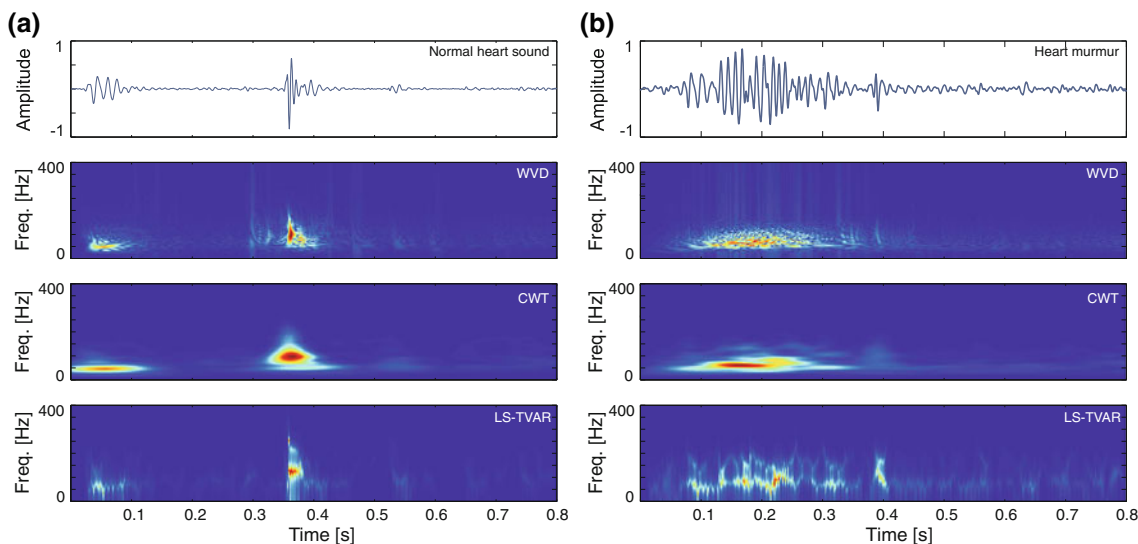

FIGURE 4. Examples of enhanced TFR for the methods of estimation considered: WVD, CWT, and LS-TVAR. (a) Normal recording; (b) PCG signal with systolic murmur.
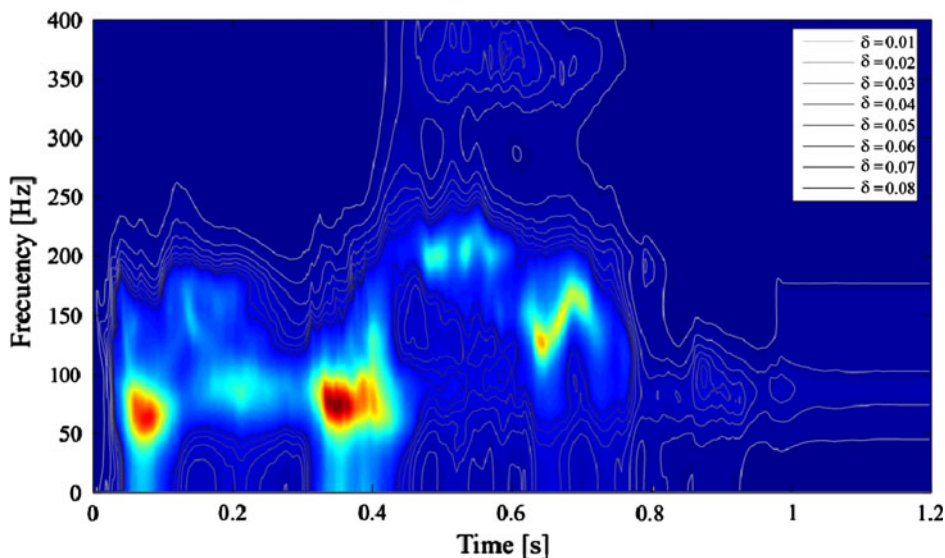

FIGURE 5. Computing the relevant area of the TFR. The different contours appraise the changes of the threshold δ.

the minimum value of the relevant variability. Figure 5 depicts different complement sets of variances depending on the value δ; the lower the threshold, the wider the relevant area.

The relevant area was empirically calculated evaluating the average variability of the classifier performance within the interval $\delta = [0.01, 0.08]$. In this article, the relevant surface of the TFR was allocated within the framework described by the time interval $0 \leq t \leq 0.8$ s and the frequency band $0 \leq f \leq 400$ Hz.

Tuning of the Tile-Based Dimensionality Reduction

On the other hand, the tiled-based approach described in the "Tiled-TFR approaches" section was carried out extracting several features from the TFR that estimates the normalized energy on each specific t–f framework. These features were used only for partitioning the TFR, but were not used for the posterior classification. For this purpose, two partition schemes were tested in the time and frequency domains: (a) First, dividing the time and frequency axes into equally spaced intervals, as suggested in Tzallas *et al.*[22]; (b) Second, performing a Quadtree partition scheme upon the TFR surface based on the pixel variance estimates, thus making smaller those t–f tiles with more information.

The Quadtree algorithm for computing the t–f framework (shown in Fig. 6a) is summarized next.

1. Using all the signals in the dataset used for training the system, compute the pixel variance matrix of every TFR surface, $\mathbf{S}_0^{(i,j)} = \text{var}\{\mathbf{S}_x(\Delta t_i, \Delta f_j)\}$, which is defined as the root of the tree decomposition.
2. Define an information threshold $\varsigma_{\min}$ and a maximum number of quad decomposition iterations N_d. Otherwise, it leads to intractable computational load.

3. Divide $\mathbf{S}_0^{(i,j)}$ into four equally sized submatrixes $\mathbf{S}_1^{(i,j)} = \text{var}\{\mathbf{S}_0(\Delta t_i, \Delta f_j)\}$ partitioning time and frequency axes in equal segments.
4. Compute the pixel variance in each submatrix $\mathbf{S}_1^{(i,j)}$. If $\text{var}\{\mathbf{S}_1^{(i,j)}\} > \varsigma_{\min}, \forall i, j$, then repeat steps 3 and 4 for further splitting of quad submatrixes, but not exceeding the number of iterations N_d. Otherwise, the values of Δt_i and Δf_j are retrieved.

The tuning of the tile size for each one of the configurations considered was carried out by estimating the average variability of the classifier performance while changing the information threshold $\varsigma_{\min}$, which ranges within the interval $\varsigma_{\min} = [0.01, 0.1]$. Specifically, Fig. 6b shows the outcomes with respect to the tile size, evidencing that a low value of $\varsigma_{\min}$ improves the accuracy of the classifier but increases the computational load.

Tuning of the Linear Decomposition Methods

The normalized matrix assessing the relevant area of the TFR is the basis for the following feature extraction. An obvious question is how to find the number of latent variables needed to obtain the best generalization for the prediction of new observations. The amount of relevant latent components, n, was chosen as follows: for PCA, this value is based on the number of breaks in the plot of the singular values, whereas for PLS, the same value is achieved by cross-validation techniques.

The n latent variables are often interpreted by plotting them in a TFR plane (as shown in Fig. 7 with $n = 4$), for systolic and diastolic murmurs, and for both decompositions considered: PCA and PLS. For PCA, considering just the first principal component it can be noticed the strong influence of the S1 event; the S2 event showed less variability. A more detailed
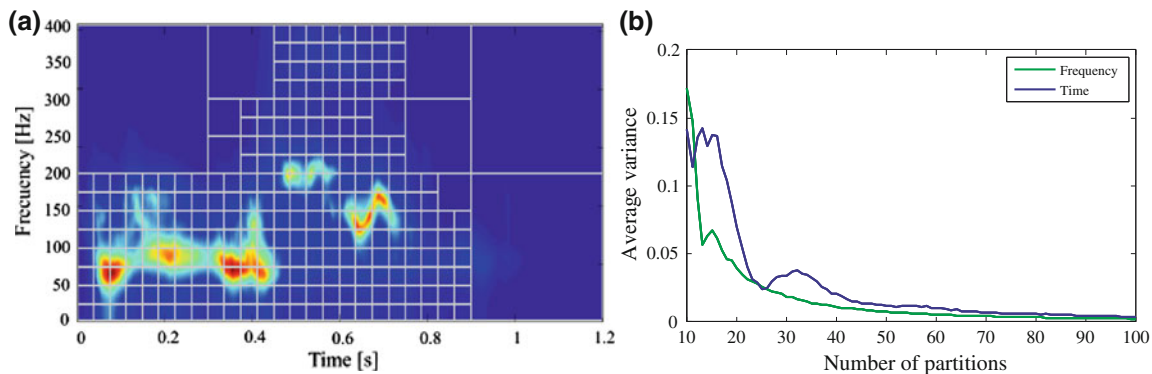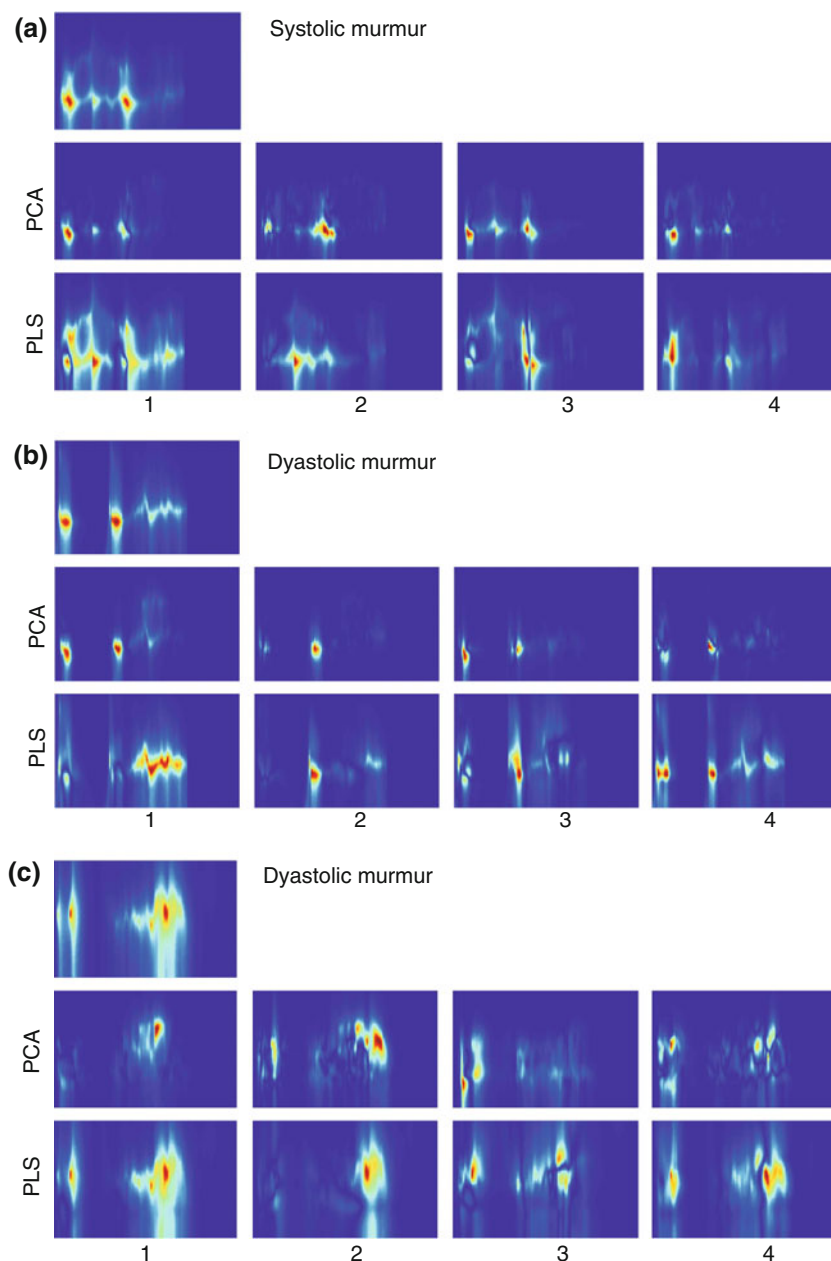

FIGURE 6. Computation of the Quadtree algorithm. (a) partition of the TFR information matrix; (b) number of partitions vs. framework variability.

FIGURE 7. First four principal vectors obtained after PCA and PLS linear decompositions multiplied by TFRs corresponding to different types of murmurs: (a) systolic murmur; (b) diastolic murmur (type 1); (c) diastolic murmur (type 2).

scrutiny of the PCA-based average reconstruction corroborates that most of the information is concentrated around the S1 and S2 events, which may be explained not only because of the instantaneous power of the PCG recordings, but also due to the irregularity of the cardiac rhythm (specially, for pathological cases). This fact becomes evident looking at the way that the intensity of S2 diminishes as n increases. At the same time, it can be noted that the influence of systole and diastole (i.e., wherever the heart murmurs are present) occurs in a different way. The presence of

systolic murmurs (see Fig. 7a) becomes more evident as the number of components increase, starting from weak values. In the case of diastolic murmurs, their presence can be clustered into two types: (a) Murmurs with weak but spread power concentrated in the lowest spectral band (e.g., mitral stenosis, aortic insufficiency, pulmonary regurgitation; noted in Fig. 7b as type 1); (b) Those with a remarkable spectral variance, and allocated in the upper part of the spectrum (e.g., aortic valve regurgitation; noted in Fig. 7c as type 2). It can be noticed that their influence vanishes as the number

of components increases. A different situation is observed for the PLS decomposition: if $n = 1$, there is a great influence of S1 and S2 events, but at the same time the presence of systolic and diastolic murmurs become quite evident.

RESULTS

In order to illustrate the problem addressed, and its difficultness, Fig. 8 shows several PCG waveforms belonging to normal and pathological states along with their respective TFR estimated by means of the LS-TVAR algorithm. Figure 8 shows that there are some normal signals whose waveform looks like pathological: waveforms x_1 and x_5 in Fig. 8a are very similar to murmurs, whereas x_1, x_3, and x_5 in Fig. 8b, at first sight, resemble normal signals.

For the sake of comparison, the feature extraction algorithms proposed were also applied to two non-parametric TFR estimations: CWT and WVD.

As commented before, the criterion used to adjust the aforementioned procedures was the maximization of the average accuracy for the detection of murmurs, given by the following definition:

$$\text{Accuracy } (\%) = \frac{N_C}{N_T} \times 100 \tag{14}$$

where N_C is the number of correctly classified patterns, and N_T is the total number of feeding patterns to the classifier. Moreover, the sensitivity and specificity measures are defined to assess the performance of the detector:

$$\text{Sensitivity } (\%) = \frac{N_{TP}}{N_{TP} + N_{FN}} \times 100;$$
$$\text{Specificity } (\%) = \frac{N_{TN}}{N_{TN} + N_{FP}} \times 100 \tag{15}$$

where N_{TP} is the number of true positives (murmurs accurately classified as murmur), N_{FN} is the number of false negatives (murmurs classified as normal signals), N_{TN} is the number of true negatives (normal signals accurately classified as normal signals), and N_{FP} is the number of false positives (normal signals classified as murmurs). The accuracy was evaluated with a simple k-nn classifier using a cross-validation strategy. The cross-validation was carried out generating 11 independent sets (i.e., folds) from the database with records randomly selected. Each fold contains the same number of records from both classes (i.e., normal and pathological). Ten sets were used for training while the remaining one was used for validation. The performance was calculated averaging the results for each fold. Moreover, the standard deviation was calculated.

After filtering and length normalization of the heart beat sounds the classification was carried out in accordance to the following training stages:

- *Tuning of the k-nn classifier*: By stepwise increasing of neighbor number, the optimal value of k is determined for the best classifier accuracy. Figure 9 illustrates the outcomes for
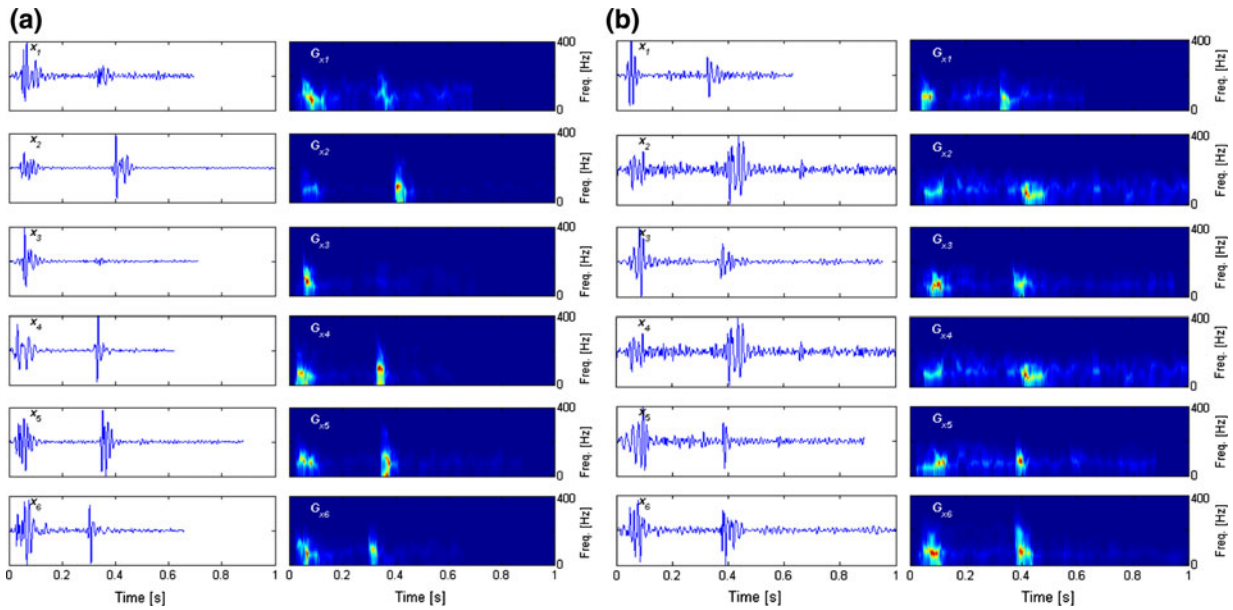
FIGURE 8. Examples of enhanced TFR for several PCG recordings. (a) normal; (b) heart murmurs.

each one of feature sets used. The figure shows that the classifier performance decreases as the number of neighbors increases. Therefore, and in order to ensure the computational stability, a value of $k = 3$ was selected.

- *Tuning of the linear decomposition methods to reduce the dimensionality.* For PCA, the information threshold and the number of components (base vectors) were the first parameters to adjust. Once again, a stepwise calculation of these parameters allowed selecting the optimum values. Figure 10a depicts the estimation of
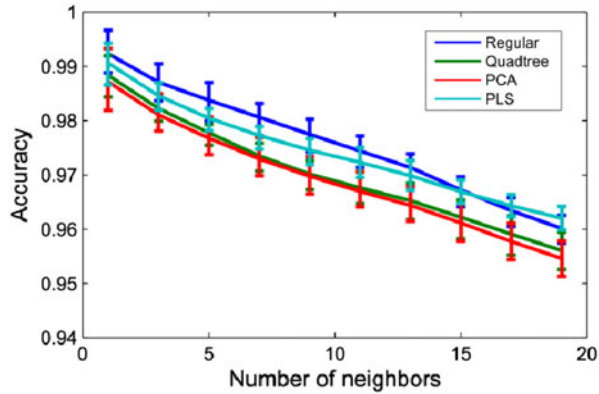
FIGURE 9. Number of neighbors vs. accuracy for the different feature extraction approaches used.

the parameter δ_{min} (see Eq. 13) using PCA, evidencing a large fluctuation of the estimates. For the PLS classifier the performance was better and more stable. Nevertheless, the value of the threshold was chosen to be $\delta_{min} = 0.07$. Figure 10b illustrates the classifier accuracy vs. the number of relevant components selected for the different schemes tested. The performance shows an asymptotic behavior starting from 8 to 10 components for both PCA and PLS techniques. However, the PLS components achieved a better performance faster than for the PCA technique. In addition, the tuning of the 2D-PCA approach was assessed by computing the accuracy for a different number of components (see Fig. 10c), which was evaluated separately for rows and columns. In the first approach, a high performance can be achieved for a low number of components ($n = 6$–8), while for the column-based approach a steady value of performance was reached just around ($n = 10$–14). During the dimensionality reduction stage, the performance was recomputed to tune the parameters of the model. Figure 10d shows that PLS and 2D-PCA provided better accuracies.

- *Tuning the tiled-based feature extraction methods.* For a regular tiling partition, the parameter to
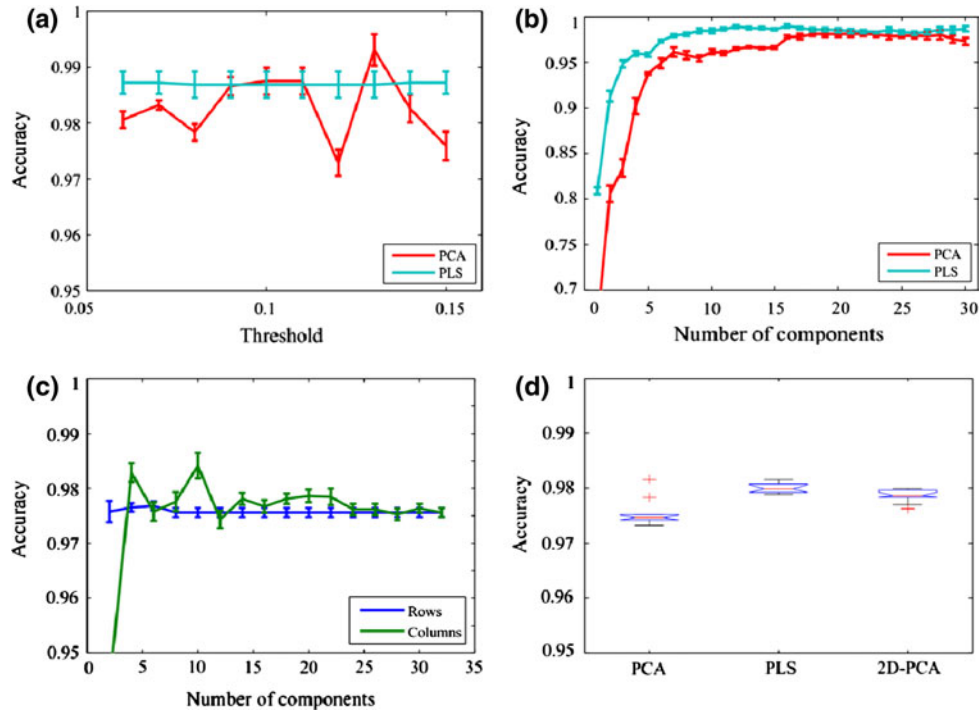

FIGURE 10. Performance of the linear decomposition approaches tested; (a) performance vs. variance threshold for PCA and PLS; (b) performance vs. number of components of PCA and PLS; (c) performance vs. number of rows and columns for the 2D-PCA approach; (d) performance of the classifier for PCA, PLS and 2D-PCA after tuning.
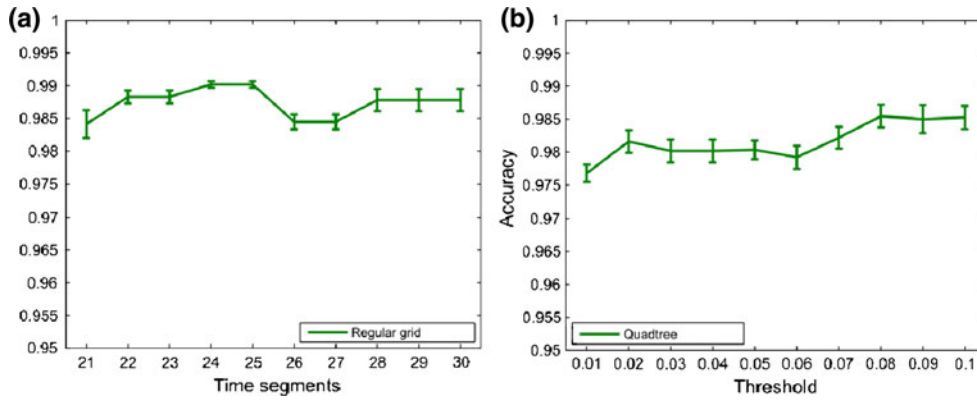
FIGURE 11. Tuning of the dimensionality reduction techniques; (a) number of time segments for the regular tiling; (b) information threshold for the Quadtree partition algorithm.
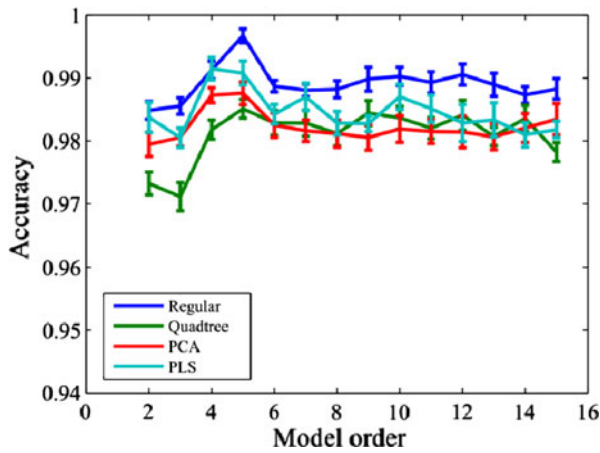

FIGURE 12. Order of the LS-TVAR model with respect to the accuracy for the different feature extraction methods used.
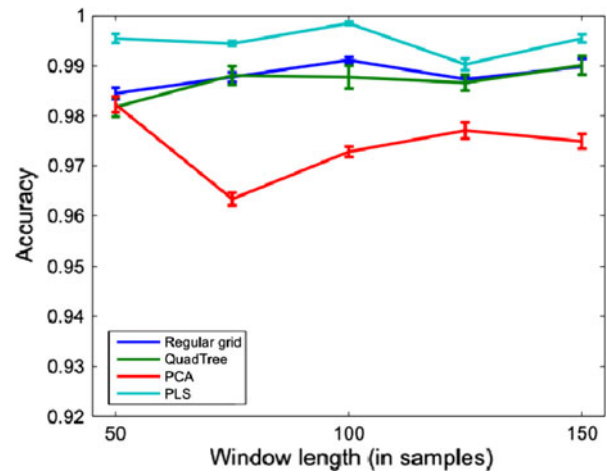

FIGURE 13. Sensibility to the window length of the LS-TVAR model to the classification accuracy.

be adjusted is the number of sub-bands (Fig. 11a); in this case, the classifier performance remains even, but the computational load increases. Regarding the Quadtree technique, the parameter to be adjusted is the threshold ς_{min} (Fig. 11b).

- *Adjusting the parameters of the TVAR model*: As shown in Fig. 12, there is a slight improvement in the accuracy as the order of the LS-TVAR model increases, but above $p = 6$–8 the performance reached a steady value. The evaluation of the accuracy with respect to the window length is shown in Fig. 13. A value of $M = 100$ can be assumed as steady. Moreover, Fig. 14 shows that the classification accuracy achieves the best performance when $\alpha = 1$, where the mean value and variance reached the best values.
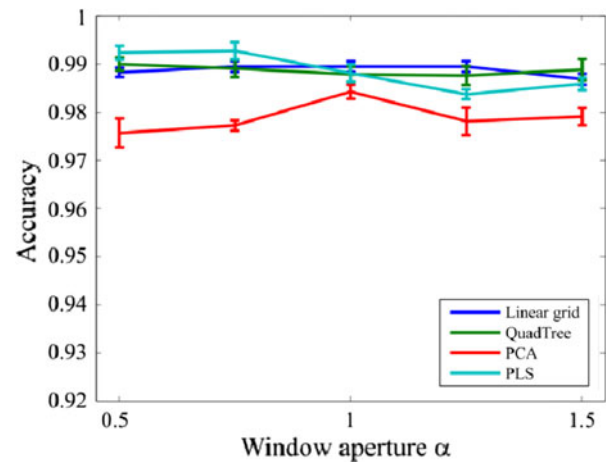

FIGURE 14. Sensibility of the classification accuracy to the aperture of the window used to estimate $\sigma_e[f]$.

In summary, the number of neighbors of the *k-nn* classifier was set to 3, and the parameters of the TVAR model were fixed to:

- The order of the TVAR model was fixed to 6.
- The window length of the LS-TVAR model was fixed to $M = 100$.
- The value of the window aperture used to calculate the variance estimator was fixed to $\alpha = 1$.

Once the dimensionality reduction was accomplished, the assemble of features used to train the system is summarized in Table 2.

These sets of features were used to train different *k-nn* classifiers with three neighbors. Table 3 summarizes the classifier performance for the different approaches tested. The accuracies were calculated averaging for each fold of the cross-validation strategy. The proposed methodology evidences a high accuracy (ranging from 97 to 99%), but with no remarkable differences among all the configurations considered. This fact is also supported in view of the estimates for sensitivity, and specificity.

The results evidence the effectiveness, stability, and dimensionality reduction capabilities of the proposed dimensionality reduction approach based on PLS applied to LS-TVAR TFR. In order to evaluate the robustness of this methodology against noise, it was tested under different conditions, adding artificial white Gaussian noise with signal-to-noise ratios (SNR) ranging from 0 to 30 dBs, with steps of 5 dBs. The results are shown in Fig. 15, where SNR = *Inf*, represents clean signals. The figure evidences a good stability of this methodology with respect to the presence of an additive white Gaussian noise: in the worst case, with SNR = 0 dB, the accuracy decreased only five absolute points.

Finally, Table 4 shows a comparison of the results achieved with the proposed methodology and other results reported in literature. The accuracy obtained with the methods described in Wang *et al.*[23] was obtained with the same database used in this study.

The results obtained using the methods in Delgado-Trejos *et al.*[8] and Quiceno-Manrique *et al.*[18] were carried out using a subset of the database used in this study that did not contain diastolic murmurs.
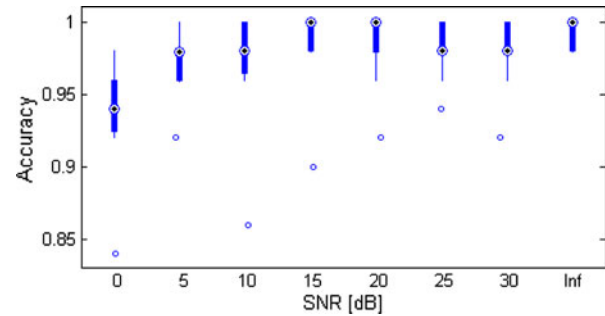

FIGURE 15. Accuracy of the PLS dimensionality reduction approach for different values of SNR.

TABLE 3. Summary of the classification performance using the proposed methodology.

Feature extraction method	Accuracy (%)	Sensitivity (%)	Specificity (%)
WVD			
Regular	98.44 ± 0.14	99.05 ± 0.08	97.84 ± 0.23
Quadtree	98.23 ± 0.07	98.55 ± 0.12	97.91 ± 0.16
PCA	97.67 ± 0.21	98.70 ± 0.11	96.66 ± 0.39
2D-PCA	98.27 ± 0.15	99.66 ± 0.19	96.91 ± 0.27
PLS	98.76 ± 0.15	99.51 ± 0.12	98.03 ± 0.20
CWT			
Regular	96.71 ± 0.36	96.80 ± 0.37	96.61 ± 0.36
Quadtree	97.04 ± 0.31	97.19 ± 0.27	96.90 ± 0.40
PCA	96.57 ± 0.28	96.46 ± 0.26	96.67 ± 0.34
2D-PCA	97.35 ± 0.32	98.39 ± 0.35	96.32 ± 0.34
PLS	97.45 ± 0.32	98.56 ± 0.44	96.38 ± 0.29
LS-TVAR			
Regular	99.00 ± 0.06	99.56 ± 0.13	98.45 ± 0.14
Quadtree	98.24 ± 0.12	98.74 ± 0.28	97.76 ± 0.13
PCA	98.17 ± 0.11	97.91 ± 0.16	98.42 ± 0.14
2D-PCA	97.56 ± 0.08	98.31 ± 0.07	96.83 ± 0.17
PLS	98.40 ± 0.19	98.31 ± 0.25	98.49 ± 0.20

TABLE 2. Summary of the features extracted for each scheme proposed.

Dimensionality reduction approach	Parameters adjusted	Value	# features (n)
Linear grid	Time partitions $N_{\Delta t}$	25	400
	Frequency partitions $N_{\Delta f}$	16	
Quadtree	Information threshold ς_{min}	0.06	355
PCA	Information threshold δ_{min}	0.07	16
PLS	Information threshold δ_{min}	0.07	20
2D-PCA	Information threshold δ_{min}	0.07	60
	Number of time components	10	
	Number of frequency components	6	

TABLE 4. Comparison with other works in the literature.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
This study	99.00	99.56	98.45
Perceptual features[23]	88.90	93.03	85.81
Time, frequency, fractal, and perceptual features[8]	96.39	95.40	95.00
Dynamic contours[18]	98.00	96.90	97.20

DISCUSSION

The classification accuracy remains stable as the order of the TVAR model changes. This means that the model order can take values lower than those suggested in the literature. This fact lead us to think that most PCG signals from the database are mono-component, which are successfully represented with a TVAR(2) model. Anyway, increasing the order complements the information given by the first two parameters modeling the noise components present in the PCG signal.

Regarding the linear decomposition, an eigenplane-based PCA technique (as simplest implementation) and a PLS technique (as more sophisticated) were considered. Likewise, the regular partition and a refined Quadtree partition were examined for the tiled-TFR approach. As shown in Table 3, the accuracy of the regular tiling outperforms the rest of techniques. However, this brute-strength approach has a higher computational load, exceeding that of the methods based on linear decomposition. Concerning to the improved object representation with PCA, 2D-PCA shows a light improvement, and reached a similar accuracy than for PLS, but with a significant reduction of the computational load. The PLS technique also evidenced a good behavior against additive white Gaussian noise.

Regarding the tuning of the feature extraction procedures it must be quoted that the computational stability of the tiled TFR-based techniques is mostly guaranteed for actual matrix sizes of HS. But for linear decomposition-based techniques, it is strongly convenient to choose a confined area of relevance on the t–f plane to achieve a stable dimension reduction, and anyhow diminishing resolution of the TFR.

For the sake of comparison, the feature extraction methodology discussed in this article was also extended to two non-parametric TFRs: WVD and CWT. A slight degradation is observed using CWT, whereas WVD shows an equivalent performance to the proposed TVAR estimation techniques (Table 3).

Furthermore, in terms of computational load and tuning of working parameters can be stated that parametric and non-parametric TFR are alike to enhance the resolution of non-stationary PCG signals.

CONCLUSIONS

In this article, the ability of an enhanced TFR and two approaches for dimensionality reduction are explored for the detection of murmurs. The enhancement of the TFR resolution is carried out by means of a parametric TVAR model-based estimation technique, allowing an improvement of the representation capabilities of the HS, and hence to a better detection of murmurs. The parametric TFR are more suitable for capturing the time-varying dynamics of the PCG signals than the common non-parametric approaches, because the resolution and distortion issues are worked out. Nonetheless, the emergent problems of this parametric approach (i.e., the selection of the model order and the estimation of the parameters) are effectively solved using locally stationary assumptions.

For the PCG signals, since the enhanced TFR tends to hold a lot of redundant data, a considerable dimensionality reduction was required. Specifically, two methods of feature extraction were analyzed: linear decomposition and tiling partition of t–f plane. The results indicate that a similar accuracy can be accomplished with both methodologies. The main difference lies on the effectiveness of the dimensionality reduction and the interpretability of the components obtained using the linear decomposition methods.

Because of the simplicity of implementation and worthy improvement of the representation capabilities determining the details of the pathological changes to identify, the suggested TFR-based feature extraction methodology appears to be very effective providing a high accuracy for the detection of murmurs. Therefore, in future studies, this framework could be applied to the recognition among different kinds of murmurs, and to the detection of pathologies for different types of non-stationary biomedical signals.

In contrast, a serious drawback of the proposed methodology is that these TFR-based feature extraction techniques require fixed length recordings (i.e., duration), so the HS must be chopped to a predetermined duration. This fact could limit the possibility to extrapolate the results to other databases, leading us to be cautious, although optimistic with the results obtained. A possibility to overcome this problem was suggested in Quiceno-Manrique et al.,[18] where dynamic time warping techniques were used to normalize the duration of each HS.

ACKNOWLEDGMENTS

REFERENCES

[1] Abramovich, Y., N. Spencer, and M. Turley. Order estimation and discrimination between stationary and time-varying (TVAR) autoregressive models. *IEEE Trans. Signal Process.* 55(6):2861–2876, 2007.

[2] Ahlstrom, C., P. Hult, P. Rask, J. Karlsson, E. Nylander, U. Dahlstrom, and P. Ask. Feature extraction for systolic heart murmur classification. *Ann. Biomed. Eng.* 34:1666–1677, 2006.

[3] Avendano-Valencia, D., F. Martinez-Tabares, D. Acosta-Medina, I. Godino-Llorente, and G. Castellanos-Dominguez. TFR-based feature extraction using PCA Approaches for discrimination of heart murmurs. *Proceedings of the 31th IEEE EMBS Annual International Conference (EMBC'09)*, 2009.

[4] Barker, M., and W. Rayens. Partial least squares for discrimination. *J. Chemomet.* 17(3):166–173, 2003.

[5] Bernat, E., W. Williams, and W. Gehring. Decomposing ERP time–frequency energy using PCA. *Clin. Neurophysiol.* 116:1314–1334, 2005.

[6] Cassidy, M., and W. Penny. Bayesian nonstationary autoregressive models for biomedical signal analysis. *IEEE Trans. Biomed. Eng.* 49(10):1142–1152, 2002.

[7] Cerutti, S., A. Bianchi, and L. Mainardi. Advanced spectral methods for detecting dynamic behaviour. *Auton. Neurosci.: Basic Clin.* 90(1):3–12, 2001.

[8] Delgado-Trejos, E., A. Quiceno-Manrique, J. Godino-Llorente, M. Blanco-Velasco, and G. Castellanos-Dominguez. Digital auscultation analysis for heart murmur detection. *Ann. Biomed. Eng.* 37(2):337–353, 2009.

[9] Deng, J., J. Yao, J. Dewald, and P. Julius. Classification of the intention to generate a shoulder versus elbow torque by means of a time frequency synthesized spatial patterns BCI algorithm. *J. Neural Eng.* 2(4):131–138, 2005.

[10] El-Segaier, M., O. Lilja, O. Lukkarinen, L. Sörnmo, R. Sepponen, and E. Pesonen. Computer-based detection and analysis of heart sound and murmur. *Ann. Biomed. Eng.* 33(7):937–942, 2005.

[11] Englehart, K., B. Hudgins, P. Parker, and M. Stevenson. Classification of the myoelectric signal using time-frequency based representations. *Med. Eng. Phys.* 21(6):431–438, 1999.

[12] Güler, I., M. Kiymik, and F. Güler. Order determination in autoregressive modeling of diastolic heart sounds. *J. Med. Syst.* 20(1):11–17, 1995.

[13] Kaipio, J., and M. Juntunen. Deterministic regression smoothness priors TVAR modelling. *Proc. IEEE ICASSP 99*, 1999, 1693–1696.

[14] Kanai, H., N. Chubachi, and Y. Koiwa. A time-varying AR modeling of heart wall vibration. In: *Proceedings on International Conference of the Acoustics, Speech, and Signal Processing, ICASSP 95*, edited by EEE Computer Society, 1995, pp. 941–944.

[15] Marchant, B. Time-frequency analysis for biosystems engineering. *Biosyst. Eng.* 85(3):261–281, 2003.

[16] Nandagopal, D., J. Mazumbar, and R. Bogner. Spectral analysis of second heart sound in children by selective linear prediction coding. *Med. Biol. Eng. Comput.* 22:229–239, 1985.

[17] Poulimenos, A., and S. Fassois. Parametric time-domain methods for non-stationary random vibration modelling and analysis—a critical survey and comparison. *Mech. Syst. Signal Process.* 20(4):763–816, 2006.

[18] Quiceno-Manrique, A. F., J. I. Godino-Llorente, M. Blanco-Velasco, and G. Castellanos-Domínguez. Selection of dynamic features based on time-frequency representations for heart murmur detection from phonocardiographic signals. *Ann. Biomed. Eng.* 38(1):118–137, 2009.

[19] Sejdic, E., I. Djurovic, and J. Jiang. Time–frequency feature representation using energy concentration: an overview of recent advances. *Digital Signal Process.* 19(1):153–183, 2009.

[20] Sullivan, G., and R. Baker. Efficient quadtree coding of images and video. *IEEE Trans. Image. Process.* 3(3):327–331, 1994.

[21] Tarvainen, M., J. Hiltunen, P. Ranta-aho, and P. Karjalainen. Estimation of nonstationary EEG with Kalman smoother approach: an application to event-related synchronization. *IEEE Trans. Biomed. Eng.* 51(3):516–524, 2004.

[22] Tzallas, A., M. Tsipouras, and D. Fotiadis. Automatic seizure detection based on time-frequency analysis and artificial neural networks. *Comput. Intell. Neurosci.* 2007:1–13, 2007.

[23] Wang, P., C. S. Lim, S. Chauhan, J. Yong, A. Foo, and V. Anantharaman. Phonocardiographic signal analysis method using a modified hidden Markov model. *Ann. Biomed. Eng.* 35(3):367–374, 2006.

[24] Yang, J., D. Zhang, A. Frangi, and J. Yang. Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 26(1):131–137, 2004.