A Stochastic Modelling Framework for the Reconstruction of Cardiovascular Signals

Diego Martín-Martínez[(1)†], Pablo Casaseca-de-la-Higuera[(1,2)*†], Marcos Martín-Fernández[(1)], Abbes Amira[(2)], Chunbo Luo[(2)], Christos Grecos[(2)], and Carlos Alberola-López[(1)]

***Abstract*— This paper presents a common stochastic modelling framework for physiological signals which allows patient simulation following a synthesis-by-analysis approach. Within this framework, we propose a general model-based methodology able to reconstruct missing or artifacted signal intervals in cardiovascular monitoring applications. The proposed model consists of independent stages which provide high flexibility to incorporate signals of different nature in terms of shape, cross-correlation and variability. The reconstruction methodology is based on model sampling and selection based on a wide range of boundary conditions, which include prior information. Results on real data show how the proposed methodology fits the particular approaches presented so far for electrocardiogram (ECG) reconstruction and how a simple extension within the framework can significantly improve their performance.**

***Index Terms*— Signal reconstruction, patient simulation, model sampling, ECG, PPG, shape model, evolution model, ARMA, PCA.**

I. INTRODUCTION

The impact of cardiovascular diseases in healthcare costs has motivated an increasing effort in the development of Wearable Health-Monitoring Systems (WHMS) by both the research community and industry. Real-time monitoring of patient's health condition can significantly decrease these costs by providing feedback information to either physicians or users so that alerting events and alarms can be triggered for health threatening awareness [1]. The accuracy of the information provided by WHMS can be seriously compromised in a connected health environment. When a patient is monitored at point of life, missing or artifacted signal records can easily appear due to, for instance, a misplaced sensor or high power noise. Data misinterpretation can be avoided by providing means of reconstructing these misleading pieces of the signal.

Fortunately, WHMS in cardiac applications bring in an intrinsic redundancy [2]. The commonly monitored variables convey mutual information due to the physiological coupling of the mechanical and electrical functions. In fact, some variables such as the Heart Rate are actually derived from the Electrocardiogram (ECG) with a high degree of accuracy. In other cases (e.g. the breathing pattern), coarse measurements can be obtained if the specific sensor is not available [3]. These cannot replace the missing ones, but provide alternative information to be used. More subtle information such as that conveyed by ECG segments can only be recovered if a prior knowledge of the ECG shape is available since no redundant information can be accounted for in this case.

A model-based synthesis-by-analysis approach can be used for reconstruction when no redundant information is available. Our recent contributions [4]–[7] proposed patient-specific stochastic models of different cardiovascular signals containing the nominal information of the behaviour of the signal as well as their variability. Specifically, the work in [6] proposed a stochastic model of the photoplethysmographic (PPG) signal able to synthesise an arbitrary number of statistically equivalent signals. The experimental evaluation showed the model capability to track physical activity, obtain statistics of clinical parameters by sampling, and reconstruct missing signal epochs. Regarding the reconstruction application, the proposal in [4] followed a similar modelling approach with the electrocardiogram (ECG) signal and showed promising results in the reconstruction of missing intervals. A substantial improvement based on jointly modelling PPG and ECG signals was recently presented in [7]. In this case, the mutual information was incorporated to the model so a more accurate signal reconstruction was achieved. A similar approach was used in [5] for the reconstruction of the respiratory waveform using the heart rate from ECG measurements.

The contributions in [4]–[7] show the utility of stochastic signal modelling of cardiovascular signals for the reconstruction of missing or artifacted signal epochs. Even though the considered signals are different in nature, the modelling and reconstruction methodologies can be defined within a common framework and thus be extended to incorporate additional signals. In this paper, we present this common framework and show how the specific proposals in [4]–[7] fit on it. Within this framework, we show how the incorporation of the respiratory signal into a joint model considering PPG and ECG improves ECG reconstruction compared with the single ECG model in [4] and the joint ECG-PPG model in [7].

The remaining of the paper is as follows: Section II

This work was partially supported by the University of Valladolid and Banco Santander under the FPI-UVa Fellowship, the Spanish Junta de Castilla y León under Grant VA136U13, the Spanish Instituto de Salud Carlos III (ISCIII) under Research Grants PI11–01492, and by the Spanish Centro para el Desarrollo Tecnológico Industrial (CDTI) under the cvREMOD (CEN–20091044) project. *Asterisk indicates corresponding author.*

[1]Laboratory of Image Processing, ETSI Telecomunicación. University of Valladolid. Spain dmarmar@lpi.tel.uva.es, marcma@tel.uva.es, caralb@tel.uva.es

[2]Audio Visual Communications and Networking Research Centre, University of the West of Scotland, Paisley, United Kingdom. pablo.casaseca@uws.ac.uk, chunbo.luo@uws.ac.uk, abbes.amira@uws.ac.uk, christos.grecos@uws.ac.uk

[†]D. Martín-Martínez and P. Casaseca-de-la-Higuera contributed equally to this work.

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

presents the common modelling and reconstruction methodology with specific examples related to our previous contributions [4]–[7]. Section III presents and discusses the reconstruction results. Finally, Section IV closes the article with the main conclusions extracted from this work.

II. METHODS

A. Model Description

The proposed model is based on the evolution of a set of parameters extracted from each heartbeat. Since all cardiovascular signals present a quasi-periodical behaviour due to the mechanics of the heart, it seems reasonable to undertake the modelling process on a heartbeat basis. A schematic of the pipeline the acquired signals undergo is presented in Figure 1. Three main stages can be identified. Specific details for each stage are provided in the following paragraphs.

1) Preprocessing: Let us consider a multichannel signal $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_K(t)]^{\mathrm{T}}$ where each channel corresponds to a different sensor in the WHMS. For the specific application we deal with in the experimental section we have $K = 3$, with $x_1(t) = x_{\mathrm{ECG}}(t)$, $x_2(t) = x_{\mathrm{PPG}}(t)$, and $x_3(t) = x_{\mathrm{RESP}}(t)$. The preprocessing stage involves a number of signal conditioning steps including (but not limited to):

- Decimation or interpolation of the signals so a common convenient sampling frequency is achieved.
- Noise filtering and detrending (e.g., baseline wandering removal in ECG).
- Heartbeat delineation.
- Time delineation of meaningful intervals when appropriate (e.g. P, Q, R, S and T waves in the ECG).

For specific details on how these steps have been performed, the reader is referred to [3]–[7]. The output of this stage is a collection of N beats $\{\mathbf{x}_n(t)\}_{n=1}^{N}$ which will directly enter the next stage.

2) Shape Modelling: The output of this stage is a parametrization of each heartbeat so that:

$$\mathbf{x}_n(t) = \mathbf{\Gamma}(\mathbf{W}[n], t) + \mathbf{r}(t), \ t \in \mathcal{T} \tag{1}$$

where $\mathbf{\Gamma}(\cdot, t) = [\Gamma_1(\cdot, t), \Gamma_2(\cdot, t), \ldots, \Gamma_K(\cdot, t)]^{\mathrm{T}}$ is a set of parametric curves, $\mathbf{W}[n] = [\mathbf{W}_1[n], \mathbf{W}_2[n], \ldots, \mathbf{W}_K[n]]^{\mathrm{T}}$ is the vector of parameters for the n-th beat whose evolution is modelled in the next stage, $\mathbf{r}(t)$ is the residue (neglected hereinafter), and $\mathcal{T}$ is the domain in which $\mathbf{x}_n$ is defined. Details for the different curves $\Gamma_k(\cdot, t)$ and their parametrizations $\mathbf{W}_k[n]$ are summarized in table I for the different signals in [4]–[7].

3) Time Evolution: The flexibility of the proposed methodology that allows considering both single- and multichannel models under a common framework relies on this evolution stage. The strategy we follow allows to separately model the evolution of each component of the parameter vector by first decorrelating the signals. To this end, we start from the parameter vector $\mathbf{W}[n]$ and transform it using Principal Component Analysis (PCA) [9] so a new parameter vector $\mathbf{Y}[n] = H \cdot \mathbf{W}[n]$ is obtained. The PCA matrix H is estimated over the whole $\mathbf{W}[n]$ vector time series since we have observed that the structure of the covariance matrix is approximately constant in time. After that, the evolution of each component can be independently modelled. Thus, the temporal evolution can be specifically defined for each of the parameters which makes the methodology flexible. Our approach so far has been using the same model for each component of $\mathbf{Y}_k$ with good results. Depending on the nature of the signal, a stationary evolution can be considered as in [5] or specific trends may be taken into account as in [4], [6], [7]. Specifically:

- The Y_k signals for the Heart Rate and Respiratory signal in [5] are directly modelled using 4th-order Autoregressive processes (AR(4)).
- The $\mathbf{Y}_k$ vectors corresponding to ECG and PPG signals are modelled in [4], [6], [7] by first splitting each component $\mathbf{Y}_{kj}$ into a trend and a residue and modelling each of them as an Autoregressive Moving Average (ARMA(2,2)) process. A schematic of this process including the decorrelation is presented in Figure 1 (inset). In order to extract the trend, a 30 samples moving average window was used –$h_{MA(30)}[n]$– as follows:

$$Y_{kj,T}[n] = Y_{kj}[n] * h_{MA(30)}[n] \quad \text{(Trend)}, \tag{2}$$
$$Y_{kj,R}[n] = Y_{kj}[n] - Y_{kj,T}[n] \quad \text{(Residue)}, \tag{3}$$

  We will follow this approach to reconstruct the ECG signal from both the PPG and respiratory signal. So, we will incorporate the respiratory signal to the joint model by sampling it synchronously with the ECG at R-peak time stamps.

TABLE I
SHAPE MODELLING INDICATIONS FOR CARDIOVASCULAR SIGNALS IN [4]–[7].

Signal	$\Gamma_k(\cdot, t)$	$\mathbf{W}_k[n]$
ECG [4], [7]	Piecewise Hermite's interpolation [8] over the meaningful parameters corresponding to the time and amplitude labels resulting from ECG delineation —onset, peak and end of all the waves (P, Q, R, S and T), plus center of each wave— (see [4] for more details on their obtention).	ℓ_k labels (see inset in Fig. 1).
PPG [4], [6]	Mixture of two Gaussians plus a linear trend warped so that the timings onset-maximum and maximum-end are always their mean values. $\gamma(t) = a_1 \exp\left[-\frac{(t-b_1)^2}{c_1^2}\right] + a_2 \exp\left[-\frac{(t-b_2)^2}{c_2^2}\right] + m_0 + m_1 t$	Timings onset-maxima and maxima-end (t_1 and t_2). Parameters of the linear trend (m_0 and m_1), and the mixture of Gaussians a_1, b_1, c_1, a_2, b_2 and c_2 parameters define two gaussian curves. All these parameters can be seen in Fig. 1 as in [4], [6].
Respiratory and Heart Rate [5]	Constant value	Actual value of the signal.

B. Signal Reconstruction from Model Sampling

The reconstruction process consists on a synthesis-by-analysis approach which generates a number of candidate

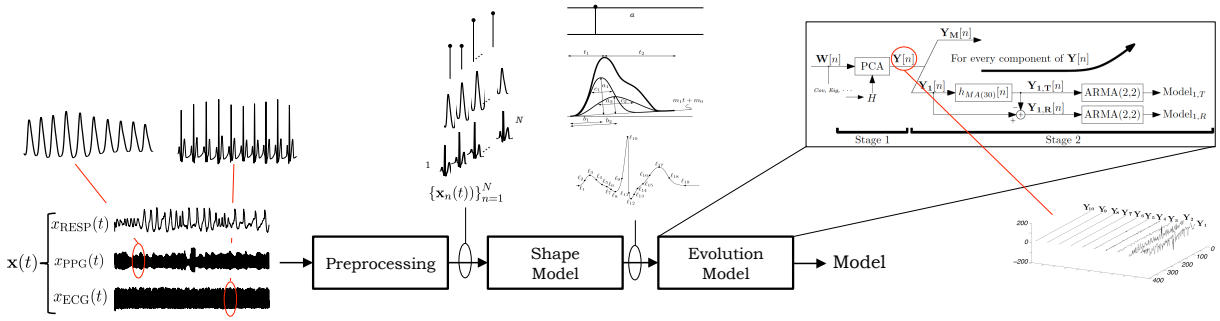

Fig. 1. General pipeline for the proposed modelling framework.

samples of the artifacted or missing epoch followed by a selection based on boundary conditions. These boundary conditions are also used to perform a refinement of the synthesized signal in joint models. The following stages can be identified:

1) Analysis of the model: The parameters of the model are estimated according to the diagram in Fig. 1 using the best (longest) available piece of the signal(s).

2) Signal synthesis by model sampling: A number of samples of the model are generated so that the number of total heartbeats exceeds the length of the epoch to be reconstructed. To this end, the temporal evolution is first simulated and correlated using the PCA matrix. Once the parameter vector is available, each heartbeat is obtained from (1).

3) Selection of the optimal sample: The optimal sample is selected based on minimizing a cost function which compares the simulated signal with a boundary condition. The reference signals are composed of one or more pieces of the available original signals . The cost function may incorporate temporal and amplitude errors between both signals (given by the parameter vector itself) as well as statistical similarity among the covariance matrices of the parameter vector. Specifically, for the works in [4], [7] and the proposal here presented:

- Temporal errors are accounted for by averaging the error between the positions of the maximum signal values for each heartbeat.
- Amplitude errors are obtained by time-averaging the point-to-point errors between the simulated signal and the boundary conditions.
- Statistical similarity is computed in terms of the William's Index [10] using the covariance matrices of the simulated data and the reference as raters and the inverse of the absolute mean error (element by element) between raters as the agreement magnitude.

The boundary conditions are picked up in a way that a few beats before and after the epoch to be reconstructed are selected together with the actual epoch if other signals are available (i.e., in joint models, see [7]).

4) Refinement of the solution: The simulation process involves the whole parameter vector even though some of the parameters might be available. If this is the case, the simulated parameters are replaced by the actual ones. Since this process will modify the covariance structure of the data, an iterative procedure is carried out over the new parameter vector to force the right correlation:

$$\mathbf{Y}_{opt}[n] = H \cdot \mathbf{W}_{opt}[n] \quad (4)$$

$$\mathbf{Y}_{opt}[n] \leftarrow GS\{\mathbf{Y}_{opt}[n]\} \quad (5)$$

$$\mathbf{W}_{opt}[n] \leftarrow H^T \cdot \mathbf{Y}_{opt}[n], \quad (6)$$

where H is the PCA matrix, and $GS\{\cdot\}$ denotes Gram-Schmidt orthogonalization over a vector series. The process is repeated until stationary convergence of the algorithm is achieved. This refinement allows including prior information in the reconstruction process thus reducing the uncertainty of the model. For instance, in [7], the PPG signal is available to reconstruct a missing ECG artifact. Since the position of the R wave can be accurately obtained from this PPG signal, the uncertainty of this parameter is removed by including it in the refinement stage.

III. Results and Discussion

To illustrate the flexibility and performance of the proposed framework, we have manually removed different length segments from a 5 min ECG registry (V_5 derivation) and reconstructed them using: 1) A single model of the ECG signal as in [4], 2) a joint ECG-PPG model as in [7], 3) a new extended joint model of the ECG, PPG and the respiratory signal. The data were obtained from a 25 years old healthy patient and were sampled at 200 Hz (ECG) and 66.67 Hz (PPG and respiratory wave), and synchronized and interpolated at 250 Hz for delineation.

Figure 2 shows the reconstruction results for a 20 heartbeats epoch. The original and reconstructed pieces as well as the point-to-point absolute error between the reconstructed and removed pieces are presented for the three models. The errors are also shown in Fig. 3 using box plot diagrams. It is worth noting that the stochastic nature of the model guarantees statistic convergence but this is not the case for point-to-point convergence. However, the flexibility of the model to incorporate boundary conditions yields accurate results in terms of point-to-point convergence, specially for joint models. The spiky nature of the error in the single ECG model reveals a variability in the position of the R peak. This variability does not affect the joint models (proposal

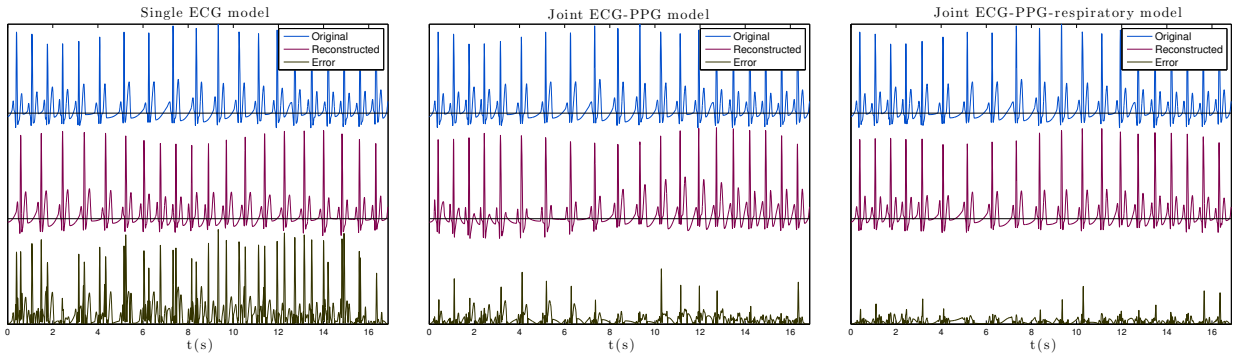

Fig. 2. Reconstruction results obtained using a single model of the ECG (left, as in [4]), a joint ECG-PPG model (top, as in [7]), and the new extended joint ECG-PPG-respiratory model (right).

here evaluated and method in [6]), since the methodology allows using the PPG signal to incorporate the actual R-peak position. In addition, the contribution of the respiratory signal to the extension here proposed improves the reconstruction results.

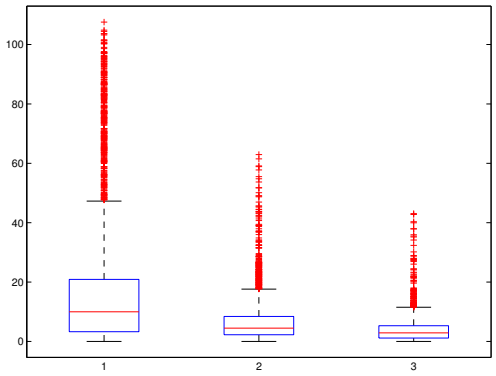

Fig. 3. Boxplots of the absolute error (point-to-point) between the reconstructed and removed pieces: 1) single model of the ECG (as in [4]), 2) a joint ECG-PPG model (as in [7]), 3) joint ECG-PPG-respiratory model.

These results show that the addition of a simple parameter can significantly improve the performance of the reconstruction method. The proposed common framework makes this extension of the model straightforward and paves the way to application-specific model definitions that can be conveniently tailored to improve the overall performance. For instance, if the accuracy of the reconstruction of the QRS intervals needs to be prioritised, the model can be defined to include additional signals conveying this information (e.g., additional ECG leads).

The PCA matrix H conveys all the mutual information among the different parameters, including cross-channel information when multiple signals are jointly modelled. Since this matrix can be of rather large size, it is important that the number of parameters to be included in the model is kept to the minimum that allows obtaining an acceptable performance for a specific application. This way, an accurate estimation of the correlation matrix can be obtained from the available samples avoiding undesirable errors in the reconstruction process.

IV. Conclusion

We have proposed a general stochastic modelling framework for the reconstruction of missing or artifacted signal periods in cardiac monitoring applications. The proposed framework is not only suited to cardiovascular signals but to other sort of signals either physiological or not. Thus, the scope of the reconstruction methodology also covers any domain in which continuous monitoring of time series is needed. Our current research is focused on extending the applicability of the modelling framework to simulate patient behavior under different conditions.

ACKNOWLEDGMENT

Authors would like to thank RGB Medical Devices LTD for kindly sharing the measurement devices.

References

[1] A. Amira, N. Ramzan, C. Grecos, Q. Wang, P. Casaseca, Z. Pervez, X. Wang, and C. Luo, *Healthcare Informatics and Analytics: Emerging Issues and Trends*. IGI Global, 2014, ch. A Reconfigurable Supporting Connected Health Environment for People with Chronic Diseases, in press.

[2] A. Mahapatro and P. M. Khilar, "Fault Diagnosis in Body Sensor Networks," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 5, pp. 252–259, 2012.

[3] L. Sörnmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, 1st ed. Burlington, MA, USA: Elsevier Academic Press, 2005.

[4] D. Martín-Martínez, P. Casaseca-de-la-Higuera, M. Martín-Fernández, and C. Alberola-López, "Cardiovascular Signal Reconstruction based on Parameterization –Shape Modeling– and Non-Stationary Temporal Modeling," in *Proceedings of the 20th European Signal Processing Conference*, Bucharest, Romania, August 2012, pp. 1826–1830.

[5] ——, "Reconstrucción de la Señal Respiratoria Basada en Modelo Conjunto con la Señal de Periodo Cardíaco," in *Proceedings of the XXX Congreso de la Sociedad Española de Ingeniería Biomédica*, San Sebastián, Spain, November 2012.

[6] ——, "Stochastic Modeling of the PPG Signal: A Synthesis–by–Analysis Approach with Applications," *IEEE. Trans. Biomed. Eng.*, vol. 60, no. 9, pp. 2432–2441, 2013.

[7] ——, "ECG Signal Reconstruction based on Stochastic Joint Modeling of the ECG and the PPG signals," in *IFMBE Proceedings*, vol. 41, Seville, Spain, 2013, pp. 989–992.

[8] G. Birkhoff, M. Schulz, and R. Varga, "Piecewise Hermite Interpolation in One and Two Variables with Applications to Partial Differential Equations," *Numer. Math.*, vol. 11, pp. 232–256, 1968.

[9] H. Abdi and L. Williams, "Principal Component Analysis," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 2, pp. 433–459, 2010.

[10] G. Williams, "Comparing the Joint Agreement of Several Raters with Another Rater," *Biometrics*, vol. 32, pp. 619–627, 1976.