

Surface ECG reconstruction from intracardiac EGM: a PCA-vectorcardiogram method

A. Kachenoura^(1,2), F. Porée^(1,2), A.I. Hernández^(1,2) and G. Carrault^(1,2)

¹ INSERM UMR 642, Laboratoire Traitement du Signal et de l'Image, Rennes, France

² Université de Rennes 1, LTSI, Campus de Beaulieu, F-35042 Rennes, France

Abstract— Cardiac implantable devices offer improved recording capabilities. Future developments have for objective to facilitate the home check-up of the patient to reduce the health care expenditure. For this purpose, a challenge deals with the patient ECG check-up from the intracardiac ElectroGraMs (EGM) delivered by the implanted prostheses. In this paper we propose a method to reconstruct the surface ECG from a set of EGM signals, based on a 3D representation of the cardiac electrical activity and Principal Component Analysis (PCA). The results, in the case of sinus rhythm, show a correlation coefficient between the real ECG and the reconstructed ECG of about 0.85.

I. INTRODUCTION

The clinical syndrome of congestive heart failure has become a major cause of morbidity and mortality. This syndrome can be treated by Cardiac Implantable Devices (CID), such as Cardiac Resynchronization Therapy Pacemaker (CRT-P) and Cardiac Resynchronization Therapy Defibrillator (CRT-D). The cardiac electrical activity acquired from the CID, named ElectroGraMs (EGM), is collected by electrodes placed on the endocardium or epicardium. Indeed, the EGM provides local information on the electric activity of a group of cardiac cells, which make their morphologies different from those observed from surface ElectroCardioGram (ECG) electrodes (considered as the reference signal for the analysis of the cardiac activity). Consequently, in order to perform a patient's check-up or modify the parameter setting of the CID, an acquisition of a standard surface ECG in an attended laboratory setting is generally required. A less expensive and less time consuming solution for performing a patient's check-up is studied in this paper. More precisely, we aim to propose a method to rebuild a surface ECG using only the intracardiac EGM recorded from CID electrodes.

II. PROBLEM STATEMENT AND REFERENCE MATERIAL

A. Signal model

The problem that we propose to study can be viewed as a direct problem (figure 1), where the outputs $\{x[m]\}_{m \in \mathbb{N}}$, representing the ECG, are considered as an unspecified nonlinear function \mathcal{F} of the inputs $\{s[m]\}_{m \in \mathbb{N}}$, representing the EGM. Then we have:

$$x[m] = \mathcal{F}(s[m]) \quad (1)$$

In the following, vectors are denoted by bold-faced lower-cases, whereas bold uppercases denote matrices. One way of approaching this problem could be to estimate a set of transfer functions directly relating each EGM signal with each ECG

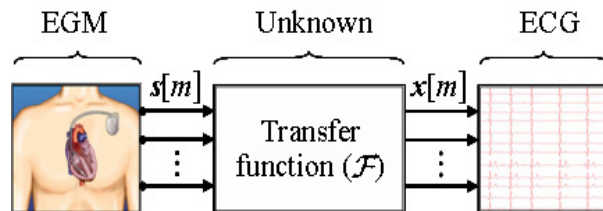


Fig. 1. Problem formalization.

signals [2]. Such an approach has, as major drawback, the fact that the analysis of the standard 12-lead ECGs can not compensate the changes in the orientation of the electrical axis caused by various extracardiac factors [5]. In this paper we propose an alternative solution, which is based on: i) the extraction of a three dimensional (3D) representation of cardiac electrical activity [6], [7] both for surface ECG (which is called VectoCardioGram, VCG) and for EGM (that we call VectoGram, VGM¹), by using the Principal Component Analysis (PCA) [1], and ii) the estimation of the transfer function between VGM and VCG. More precisely, a two-step procedure is proposed to rebuild a surface ECG. It consists in a training step, which aim to identify the transfer function between the EGM and the ECG and the reconstruction step which is devoted to the estimation of surface ECG by exploiting only the EGM, and the estimated transfer function in the training step.

Nevertheless, before detailing our procedure, let us give a brief description both of the 3D representation of cardiac electrical activity and PCA and justify why these two tools are used in our approach.

B. The 3D representation of cardiac electrical activity

The VCG is the methodological extension of ECG that provides a 3D representation of the cardiac electrical field. More precisely, the VCG is an orthogonal lead system that reflects the electrical activity in the three perpendicular directions X , Y , and Z . For such a system, the interpretation is not confined to findings in individual leads (case of ECG), but additional information is acquired through the visualisation of 3D loops together with its projection onto the YZ (frontal), XZ (sagittal) and XY (transverse) planes [6], [7]. Although

¹The VGM concept has never been introduced in the literature before us.

the 12-lead ECG is considered as the reference signal for the analysis of the cardiac activity due to the existence of a number of rules for its interpretation, the VCG contains some useful information for certain applications. Indeed, it is well known that the VCG is superior to the ECG in showing phase differences between electric events in different parts of the heart. In addition, contrary to the standard 12-lead ECG, the analysis based on VCG loops has been found to: i) compensate the changes in the electrical axis caused by various extracardiac factors [5] and ii) give a solution to the time synchronization problem which arises in cardiac data.

C. The principal component analysis

PCA, which is closely related to Karhunen-Loève Transform (KLT) (also known as Hotelling transform), is a classical technique in statistical data analysis, features extraction and data compression. The purpose of PCA is to derive a relatively small number of uncorrelated linear combinations (principal components) of a set of random zero-mean variables while retaining as much of the information from the original variables as possible. Typically, the PCA of vector process $\{\mathbf{x}[m]\}_{m \in \mathbb{N}}$ consists in looking for an $N \times P$ orthonormal linear transform \mathbf{W} (P smaller or equal to N), such that:

$$\mathbf{z}[m] = \mathbf{W}^T \mathbf{x}[m] \quad (2)$$

where the components of the vector process $\{\mathbf{z}[m]\}_{m \in \mathbb{N}}$ are mutually uncorrelated. Sometimes we need the column of the matrix \mathbf{W} ; if we denote w_{np} ($1 \leq n \leq N$ and $1 \leq p \leq P$) the elements of the \mathbf{W} the model of (2) can also be written as:

$$z_p[m] = \sum_{n=1}^N w_{np} x_n[m] = \mathbf{w}_p^T \mathbf{x}[m] \quad (3)$$

It is straightforward to show that the PCA problem can be converted to the eigenvalue problem of the covariance matrix \mathbf{R}_x [1]. Thus, if we denote $[e_1 \dots e_P]^T$ the eigenvectors of \mathbf{R}_x corresponding to the eigenvalues $(\lambda_1, \dots, \lambda_P)$ where $\lambda_1 \geq \dots \geq \lambda_P$, the first principal component of $\mathbf{x}[m]$ is $z_1[m] = e_1^T \mathbf{x}[m]$. Likewise, the P -th principal component is obtained by $z_P[m] = e_P^T \mathbf{x}[m]$.

III. A TWO-STEP PROCESSING APPROACH

The aim of this section is to provide more information (see figure 5) about the two steps of our procedure, namely the training step and the reconstruction step.

A. Training step

This step (figure 2(a)) can be decomposed into three sub-steps: the identification of the orthonormal linear transform \mathbf{W} , between the ECG ($\mathbf{x}[m]$) and the VCG ($\mathbf{z}_{VCG}[m]$), the 3D registration of the VGM loops ($\mathbf{z}_{VGM}[m]$) and the estimation of the filter $\mathbf{h}[m]$ between the VCG and the registered VGM loops.

Identification of \mathbf{W} . Let us assume that $\mathbf{x}[m] = [x_1[m], \dots, x_N[m]]^T$ (where $N \geq 3$ and $m = 1, \dots, M$) represents the ECG of L successive heartbeats. The orthonormal

linear transform \mathbf{W} is estimated by applying the PCA on $\mathbf{x}[m]$ so that the following result holds:

$$\mathbf{z}_{VCG}[m] = \mathbf{W}^T \mathbf{x}[m] \quad (4)$$

It is worth noting that once the $N \times P$ matrix \mathbf{W} is identified, only the components of interest are considered. In our case, we just take into account the three principal components, corresponding to the three largest eigenvalues of the covariance matrix \mathbf{R}_x , which provides us a $N \times 3$ matrix \mathbf{W} .

The VGM 3D registration. In order to improve the precision of some automatic ECG analysis algorithms, L. Sörnmo proposed in [5] a pre-processing method to compensate for heartbeat morphology variations during an ECG/VCG recording. This method, based on four steps (translation, scaling, rotation and time synchronisation) is applied here to the registration of VGM loops. Let us assume that the VCG loop is considered in our case as a reference loop. As for ECG, let $s[m] = [s_1[m], \dots, s_K[m]]^T$ (where $K \geq 3$ and $m = 1, \dots, M$) be the EGM of L successive heartbeats. The VGM, $\mathbf{z}_{VGM}[m]$, is firstly derived by applying the PCA on $s[m]$, where we just take into account the three largest eigenvalues of the covariance matrix \mathbf{R}_s . Then, both $\mathbf{z}_{VCG}[m]$ and $\mathbf{z}_{VGM}[m]$ are segmented into L non-overlapping blocks of equal length T , $\mathbf{z}_{VCG\ell}[m]_{\ell=1, \dots, L}$ and $\mathbf{z}_{VGM\ell}[m]_{\ell=1, \dots, L}$, respectively. Now let us consider the expectations of $\mathbf{z}_{VCG\ell}[m]_{\ell=1, \dots, L}$ and $\mathbf{z}_{VGM\ell}[m]_{\ell=1, \dots, L}$ estimated by averaging over the number of heartbeats L using the following sample formula [3]:

$$\bar{\mathbf{z}}_{VCG}[m] = \frac{1}{L} \sum_{\ell=1}^L \mathbf{z}_{VCG\ell}[m] \quad (5)$$

$$\bar{\mathbf{z}}_{VGM}[m] = \frac{1}{L} \sum_{\ell=1}^L \mathbf{z}_{VGM\ell}[m] \quad (6)$$

In order to include the time synchronisation step, it is necessary to subtract 2Δ samples to the reference loop $\bar{\mathbf{z}}_{VCG}[m]$, which provides us a new $3 \times T - 2\Delta$ reference loop $\tilde{\mathbf{z}}_{VCG}[m]$. The estimation of the rotation \mathbf{R} , the scaling α and the shift time τ is based on a model in which the VGM loop $\bar{\mathbf{z}}_{VGM}[m]$ is related to the reference loop $\tilde{\mathbf{z}}_{VCG}[m]$ as follows:

$$\bar{\mathbf{z}}_{VGM}[m] = \alpha \mathbf{R} \tilde{\mathbf{z}}_{VCG}[m] \mathbf{J}_\tau \quad (7)$$

where $\tau = -\Delta, \dots, \Delta$ and $\mathbf{J}_\tau = [\mathbf{0}_{\Delta-\tau} \mathbf{I} \mathbf{0}_{\Delta+\tau}]^T$. The dimensions of the left and right zero matrices are equal to $\Delta - \tau \times T - 2\Delta$ and $\Delta + \tau \times T - 2\Delta$, respectively.

It is easy to show that the estimation of \mathbf{R} , α and τ can be reduced to the following minimization problem [5]:

$$\xi_{\min}^2 = \min_{\alpha, \mathbf{R}, \tau} \|\tilde{\mathbf{z}}_{VCG}[m] - \alpha \mathbf{R} \tilde{\mathbf{z}}_{VGM}[m] \mathbf{J}_\tau\|_F^2 \quad (8)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. More precisely, the minimization of the previous equation is performed by first finding the estimates $\hat{\alpha}$ and $\hat{\mathbf{R}}$ by fixing τ . The optimal estimates α , \mathbf{R} and τ are then determined by evaluating the error ξ^2 for different values of τ (see [5] for more details).

Estimation of the filter $h[m]$. In order to estimate $h[m]$, we must estimate three different transfert functions $h_1[m]$, $h_2[m]$ and $h_3[m]$ between each row of the output vector $\tilde{z}_{VCG}[m]$ and the input vector $\tilde{z}_{VGM_R}[m]$. For the sake of readability let us consider the estimation of $h_1[m]$, which is characterized by the input-output relationship:

$$\tilde{z}_{VCG_1}[m] = (\tilde{z}_{VGM_{R_1}} * h_1)[m] = \sum_{i=0}^{Lh-1} h_1[i] \tilde{z}_{VGM_{R_1}}[m-i] \quad (9)$$

where $*$ denotes the convolution operator. In vector notation it reads:

$$\tilde{z}_{VCG_1}[m] = \mathbf{h}_1 \tilde{z}_{VGM_{R_1}}[m] \quad (10)$$

where the Lh -dimensional parameter vector $\mathbf{h}_1 \stackrel{\text{def}}{=} [h_1(0), \dots, h_1(Lh-1)]^T$ is the impulse response of a Linear Time Invariant (LTI) filter, and $\tilde{z}_{VGM_{R_1}}[m] \stackrel{\text{def}}{=} [\tilde{z}_{VGM_{R_1}}[m], \dots, \tilde{z}_{VGM_{R_1}}[m-Lh+1]]^T$. An estimate, $\hat{\mathbf{h}}_1$, of \mathbf{h}_1 , can thus be derived from the general Wiener-Hopf equation [4] which relates the optimal LTI filter to the covariance matrix of the output $\mathbf{R}_{\tilde{z}}$ and to the intercorrelation between the output and the input $\mathbf{R}_{\tilde{z}\tilde{z}}$, such that:

$$\hat{\mathbf{h}}_1 = \mathbf{R}_{\tilde{z}}^{-1} \mathbf{R}_{\tilde{z}\tilde{z}} \quad (11)$$

The same result can be directly derived for \mathbf{h}_2 and \mathbf{h}_3 .

B. Reconstruction step

This step (figure 2(b)) is devoted to the estimation of surface ECG, by exploiting the EGM and different parameters identified in the training step. To do so, we suppose that we only observe the EGM of Q successive heartbeats, denoted by $\mathbf{s}[m] = [s_1[m], \dots, s_K[m]]^T$ (where $K \geq 3$ and $m = 1, \dots, M'$). Then, $\mathbf{s}[m]$ is segmented into Q non-overlapping blocks of equal length T , $\mathbf{s}_q[m]_{q=1, \dots, Q}$ and the PCA is applied on each block, which provides us Q VGM blocks $\mathbf{z}_{VGM_q}[m]_{q=1, \dots, Q}$ (as for the training step, only the three first principal components are taken into account). Finally, the Q heartbeats surface ECG are estimated one by one as following:

$$\forall q \in \{1, \dots, Q\}, \hat{\mathbf{s}}_q[m] = (\hat{\alpha}_\tau \hat{\mathbf{R}}_\tau \mathbf{z}_{VGM_q} \mathbf{J}_\tau^T * \mathbf{h})[m] \mathbf{W}^T \quad (12)$$

IV. RESULTS

A data set issued from 10 patients with sinus rhythm (P1 to P10) is used for evaluating the proposed method. Each record of the database is composed of 12 surface ECG channels, namely I, II, III, AVR, AVL, AVF, V1, to V6, and 3 to 11 EGM leads (11 EGM for P1, P2, P3, P4, six EGM for P5, P6, five EGM for P7, P8, P9 and 3 EGM for P10) depending on the CID type. The ECG and EGM are simultaneously recorded with a GE Cardiolab station during the implant of CIDs with a sampling rate equal to 1000 Hz. Each patient's data is segmented into two blocks. The first block contains L heartbeats of ECG and EGM is used in the training step in order to estimate the transfer function between EGM and ECG (see section III-A), whereas the second block containing Q

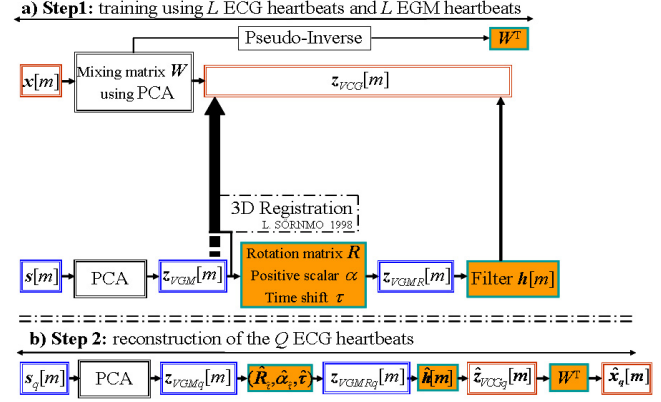


Fig. 2. Detailed representation of our procedure: a) training step, b) reconstruction step.

heartbeats is devoted to the reconstruction step (see section III-B). In order to evaluate the quality of the reconstruction the correlation between the real ECG and the reconstructed signal is processed.

Figures 3(a) and 3(b) and 3(c) show an example of 11 EGM channels recorded from the CID and the real surface 12-leads ECG and the reconstructed ECG (ECGR), respectively. Clearly, the 12 ECG channels are well estimated (figure 3(c)), whereas the original EGM (figure 3(a)) input signals are really morphologically different from the real ECG. Indeed, figure 3(d) shows, for each heartbeat and each ECG channel, that the reconstruction errors are practically insignificant. This result is also confirmed by the correlation coefficient (figure 4), which is about 0.90 for the three distinct heartbeats.

Figure 5, shows the comparative performance obtained by applying our procedure by exploiting: i) all the available EGM (black bars) and ii) only three EGM channel (gray bars). It is very interesting to note that the performance obtained using three EGM are quasi-identical than those obtained by all available EGM, namely 11 EGM in the case of P1, P2, P3 and P4, six EGM in the case of P5 and P6 and five EGM for P7, P8 and P9. In addition, when considering the total of the database, the global correlation for each ECG channel is about 0.85.

V. CONCLUSION

We propose in this paper a procedure in order to facilitate the home check-up of patients by using the intracardiac EGM derived by the implantable prostheses. More precisely our procedure aims at reconstructing the surface 12-lead ECGs using the 3D representation of cardiac electrical activity and PCA. The results obtained on patients with sinus rhythm show a correlation coefficient between the real ECG and the reconstructed ECG about 0.85. In addition, our procedure seems to provide quasi-identical performance by using only three EGM or a high number of EGM (five, six and 11 EGM channels in our case). This last result is very useful in practical

case. Indeed, most of CIDs provide only three implantable electrodes. Current work is directed to the clinical evaluation of more patients, especially for those presenting an abnormal heart rhythm such as the presence of premature ventricular beats.

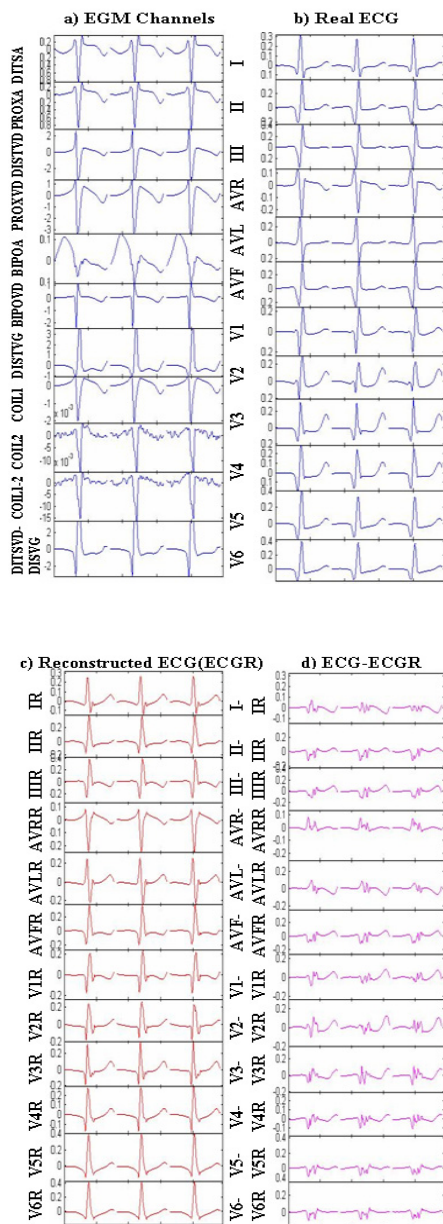


Fig. 3. Example of ECG reconstruction from EGM for P2: a) EGM, b) real ECG, c) reconstructed ECG (ECGR) and d) reconstruction error.

REFERENCES

[1] K. L. DIAMANTARAS and S.Y. KUNG, *Principal Component Neural Networks: Theory and Applications*, Wiley-Interscience Publication, New York, February 1996.
 [2] M. GENTIL, F. POREE, A. I. HERNANDEZ and G. CARRAULT, "Surface electrocardiogram reconstruction from cardiac prosthesis electrograms," *EMBE'05, European Medical and Biological Engineering Conference*, pp. 2028F1-6, Prague, Czech Republic, November 2005.

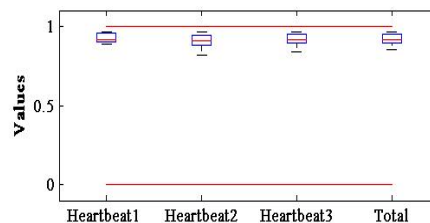


Fig. 4. The correlation coefficient, obtained for P2, between ECG and ECGR for each heartbeat.

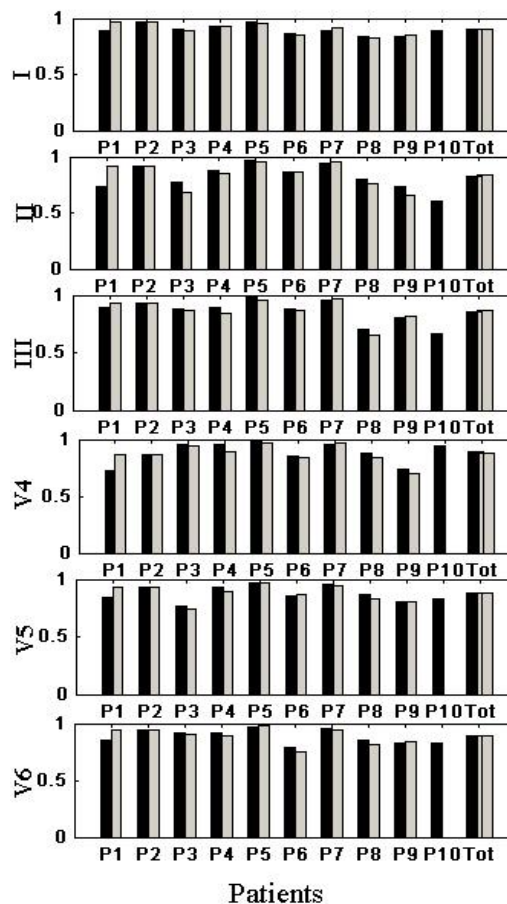


Fig. 5. The correlation coefficient between ECG and ECGR for each patient: using all EGM channels (black bars), using three EGM channels (gray bars).

[3] P. McCULLAGH, *Tensor Methods in Statistics*. Chapman and Hall, Monographs on Statistics and Applied Probability, London, 1987.
 [4] B. PICINBONO, *Random Signal and Systems*. Prentice-Hall, New Jersey, 1993.
 [5] L. SÖRNMO, "Vectorcardiographic Loop Alignment and Morphologic Beat-to-Beat Variability," *IEEE Transaction on Biomedical Engineering*, vol. 45, no. 12, pp. 1401–1413, December 2004.
 [6] L. SÖRNMO and P. LAGUNA, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*, Elsevier, Academic Press, USA, July 2005.
 [7] T. WINSOR, *Primer of Vectorcardiography*, Lea and Febiger, Philadelphia, 1972.