Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Independent Component Analysis of Parameterized ECG Signals

Jarno M. A. Tanskanen, *Member, IEEE*, Jari J. Viik, and Jari A. K. Hyttinen, *Member, IEEE*

***Abstract*—Independent component analysis (ICA) of measured signals yields the independent sources, given certain fulfilled requirements. Properly parameterized signals provide a better view to the considered system aspects, while reducing the amount of data. It is little acknowledged that appropriately parameterized signals may be subjected to ICA, yielding independent components (ICs) displaying more clearly the investigated properties of the sources. In this paper, we propose ICA of parameterized signals, and demonstrate the concept with ICA of ST and R parameterizations of electrocardiogram (ECG) signals from ECG exercise test measurements from two coronary artery disease (CAD) patients.**

I. Introduction

ICA [1] is a widely used signal processing approach. The ICA signal mixing model is given by

$$\mathbf{Y} = \mathbf{AX}, \tag{1}$$

where ICs, i.e., sources in the rows of **X** are mixed at measurement sites according to a mixing matrix **A**, yielding measurements in the rows of **Y**. For ICA, the sources are to be independent and nongaussian [1]. Thereafter, the ICs may be obtained by a numerical, iterative algorithm, such as FastICA [2,3]. From ICA, ICs appear in **X** with indeterminate energies and signs, and in an unknown order.

Our aim is that the ICs of the parameterized signals would convey information in a form more readily interpretable by a human or a computational algorithm for the problem at hand. For many types of real-world signals, e.g., ECGs and other biomedical signals, fulfillment of ICA requirements are usually not known. Nevertheless, the ICA mixing model (1) may be assumed valid in many cases, and ICA may be attempted. ICA has been widely applied in numerous applications. Applications related to our work are, for example, fetal ECG separation [4] and atrial activity estimation [5]. While ICA has been applied in signal parameter, or feature, extraction [1], it is little acknowledged that it may also be possible and advantageous to calculate ICA on signal parameters themselves.

Bioelectromagnetic processes in the heart manifest themselves in ECG measurements. Each ECG electrode measures a different view to the bioelectromagnetic activity of the heart. In this work, we assume that the bioelectromagnetic processes are mixed at the ECG electrodes according to (1). In the human body, there most likely exist local noise sources, which are not measured by all the electrodes. This would correspond to having zeros in **A** (1). Here, we assume that such noise sources are negligible for the analysis, and appear as noise in the found ICs in an indeterminable fashion. On the other hand, however, if the actual number of independent sources, including both noise and payload signal sources, is less than the number of the measurements, ICA will yield only the true number of ICs.

To demonstrate ICA of parameterized ECG signals, we apply ICA on two parameterizations of ECG exercise test measurements from two CAD patients, drawn randomly from the ECG exercise test data base of Tampere University Hospital, Finland.

II. Signal Parameterization for ICA

Based on trivial properties of convolution, some common signal parameters also fulfill the ICA mixing model, given that the original measured signals comply with it. By signal parameterization we mean construction of a new signal from any local or global properties, i.e., parameters, of the original signal. These properties may be related to a priori known features of the sources. In some cases, the appropriate parameterization may be such that it greatly decrease the amount of data, thus lowering the computational burden of the subsequent analysis. ICs of properly parameterized signals may also be more readily interpretable than the measurements themselves, or their ICs.

As readily stated by the mixing model (1), each and every measured sample is a linear combination of the samples of the source signals at the same time. Therefore, a set of signals, in which each signal consists of samples from the corresponding original measurement at the time points **S**, i.e.,

$$\mathbf{Y'} = \mathbf{Y}(\mathbf{S}) \tag{2}$$

also fulfills the ICA mixing model (1). The set of ICA input signals is now the set of parameterized signals composed of all the original measurements sampled at times **S**. For example, **S** may consist of the time points of local minima or maxima of a measured signal.

Also, it can be shown that signals constructed from time averages of the original measurement samples, or signals resulting from FIR filtering, fulfill the ICA signal model, and may thus be subjected to ICA. However, it is to be noted that even if the ICA algorithm converged and produced ICs, av-

The work of J. M. A. Tanskanen was funded by funded by Academy of Finland grant number 206996.

J. M. A. Tanskanen (phone: +358 50 387 1347; fax: +358 3 3116 4013; e-mail: tanskanen@ieee.org), J. J. Viik (e-mail: jari.viik@tut.fi), and J. A. K. Hyttinen (e-mail: jari.hyttinen@tut.fi) are with Ragnar Granit Institute, Tampere University of Technology, P.O. Box 692, FI-33101 Tampere, Finland.

eraging will make the components more Gaussian. Thus, one must pay special attention to the proper interpretability of the components, and avoid long averaging windows.

III. ICA OF PARAMETERIZED ECG SIGNALS

For demonstration purposes, we use the six signals from a 12-lead ECG, measured on the patients chest around the heart, i.e., the signals V_1, V_2, …, V_6 [6]. These six signals may best confirm to the ICA signal model; it can be expected that these signals carry contributions from the independent bioelectromagnetic phenomena of the heart with fewer disturbances than, for example, the measurements from the limb electrodes, i.e., that source signals appeared at the measurement sites differently scaled but otherwise undistorted. The ECG parameters were calculated using ECG Exercise Workstation Case8000ws software from General Electric Medical Systems Information Technologies. It is desired that the ICs of the parameterized signals would display diagnostic information carried by the parameters more clearly than what is observable from the original measurements or their ICs. Here, we demonstrate the proposed method with ECG signal parameterizations based on local maxima and averages.

There exists numerous ECG parameterizations; our database contains 16 different parameterization of the measured exercise test ECGs, and it is not yet exactly known, which of them would best lend themselves to the proposed ICA analysis. R and ST parameterizations are two potential choices. It is to be noted that in order for the ICA mixing model (1) to be valid, the parameters have to be derived from signal amplitudes; ECG parameters describing time durations, e.g., the time periods between consecutive R waves, do not comply with the ICA mixing model. Even with proper parameterizations, it may sometimes be hard or impossible to make an ICA algorithm converge, because of missing or otherwise bad data due to, for example, bad electrode contacts. It is also otherwise possible that the ICA does not converge, or that several runs of ICA are needed.

A. Exemplary ECG Parameterizations and Their ICAs

Two seconds of exemplary ECG signals during the exercise phase of a 27-minute exercise test are shown in Fig. 1A, and the corresponding ICs in Fig. 1B. An exercise test consists of three phases; initial rest, exercise, and recovery. In this paper, we have considered the ECG signals and the parameterizations from the start of the exercise phase until the end of the recovery phase. In Fig. 1A, the first R wave and an approximate ST-segment duration of V_6 are indicated. Let us denote the maximum amplitude of an R wave by R. R and ST parameterizations of the ECG, which is seen in part in Fig. 1A, are seen in Figs. 2A and B, respectively for the duration of the exercise and recovery phases of the exercise test, along with the results of ICA calculated on the parameterized signals. R and ST parameterizations along with the corresponding ICA results from another CAD patient are show in Figs. 2C and D.

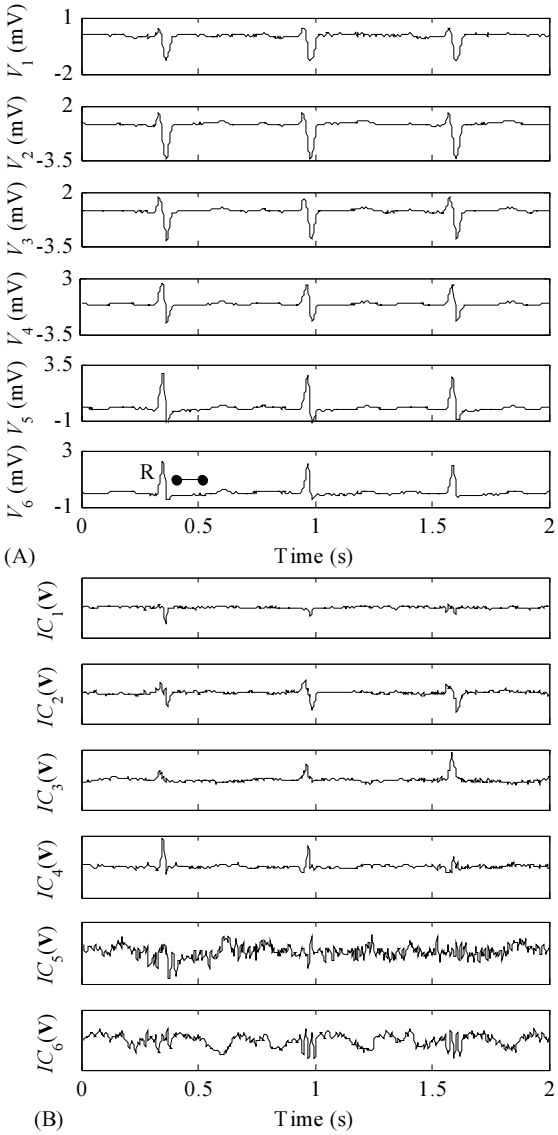

Fig. 1. (A) Two seconds of ECG signals V_1 through V_6 of a 12-lead ECG exercise test measurement from a CAD patient during a latter part of the exercise phase. An exemplary R wave is indicated for signal V_6. The corresponding R value is the maximum signal amplitude of the wave to the right of the "R". An approximate ST-segment extent for one heart beat complex is indicated with the line with the dots at its ends. *ST60* is the signal amplitude 60 ms after the start of the ST-segment, and our *ST* parameter value is the average of three consecutive *ST60* values. (B) ICs from ICA on the signals in (A). Note that the amplitudes and signs of the ICs, and the IC order, are arbitrary.

ST-segment level is traditionally given by the signal amplitude, for example, 60 ms after the start of the ST-segment and denoted by *ST60*. Here, we have calculated each *ST* value as an average of three consecutive *ST60* values. Although there exist more powerful noise reduction methods, averaging consecutive *ST60* values over a number of heart beat complexes is commonly used in the clinic. In Figs. 1B and 2, the matrixes composed of the signals V_m, R_{Vm}, and ST_{Vm}, m = 1, …, 6, are denoted by **V**, **R** and **ST**, respectively. IC_m(**R**), m = 1, …, 6, for example, denotes the mth IC from the ICA calculation with **V** as the input.

R reflects the intensity and direction of progressing ven-

tricular heart muscle depolarization, i.e., contraction. The different simultaneously measured parameterized signals R_{Vm} (Figs. 2A and C) reflect this activity as seen from different angles with respect to the heart, and their ICs present us with information on the independent processes forming the **R** as the exercise test progresses. *ST60* carries information on the slow phase of ventricular repolarization. Its variation from the ECG baseline indicates CAD [7]. This information is captured in **ST** (Figs. 2B and D), and is routinely used in CAD diagnostics [8].

B. Results of ICA

ICs from ICA on the original ECG signals in Fig. 1A are shown in Fig. 1B, and they display the independent electrical processes of the heart, some of which are not clearly identifiable from the original ECG signals, such as the activity seen in signals IC_5(**V**) and IC_6(**V**). Especially, the high frequency oscillations of IC_6(**V**) in Fig. 1B might have diagnostic value. Also, comparing Figs. 1A and B, the R waves and the following downward waves are seen to be divided differently into four uppermost ICs in Fig. 1B.

The R parameterization of the ECG seen in part in Fig. 1A, are shown in Fig. 2A for the exercise and recovery periods, lasting altogether approximately 15 min, along with the corresponding ICs. The vertical dashed line indicates the end of the exercise phase, which is followed by the recovery phase. A short section of bad ECG measurements is seen in signals R_{V4}, and R_{V5}, as an immediate drop of signal level around the sample number 200 in Fig. 2A. Contributions of this incident can be seen in all ICs in Fig. 2A. At least IC_5(**R**) and IC_6(**R**), and perhaps also IC_4(**V**), are seen to respond to the end of exercise, in Fig. 2A.

In the R parameterizations of the ECG from the same patient (Fig. 2B), ST_{V4}, ST_{V5}, and ST_{V6}, are clearly seen to exhibit ST level depression typical to CAD, which is mostly captured by IC_2(**ST**) and IC_4(**ST**). The peak following the end of exercise is mostly captured in IC_1(**ST**). During the exercise in Fig. 2B, IC_3(**ST**) displays two most curious peaks, which can barely be identified in the parameter signals ST_{V1} and ST_{V2}, but most probably not from the original ECG. The fact that the two notches in ST_{V1} and ST_{V2} turn out to be part of the same independent signal, which is otherwise mostly noise, cannot be deduced from the parameter signals. Thereafter, in Fig. 2B, IC_5(**ST**) and IC_6(**ST**) carry mostly noise.

Considering the R parameterization of the other CAD patient's ECG, seen in Fig. 2C, the most obvious phenomenon is the signal level drop after the end of exercise observable in every R parameterized signal, which recovers after a short while. This drop is seen to be captured in different fashions in IC_2(**R**), IC_3(**R**), and IC_4(**R**), and the rising edge perhaps also in IC_1(**R**). In Fig. 2D, only ST_{V1} is seen to exhibit the CAD related ST level depression, which is probably captured in IC_3(**ST**) and IC_6(**ST**), whose shape may be of interest. Finally, the peak seen in ST_{V2} though ST_{V6}, in Fig. 2D, is mostly due to a single independent process IC_1(**ST**).

IV. Conclusions and Discussion

In this paper, we have proposed and demonstrated ICA of parameterized ECG signals. As appropriate signal parameters reveal desired signal properties, ICs of thus parameterized ECG signals may as well display the desired aspects of the question at hand much more clearly than the ICs of the original signals, while possibly also reducing the amount of data to a fraction of the original. The ICs of the parameterized ECG measurements clearly display processes of the heart, which are not immediately visible in the traditional ECG assessment, or in the ICs of the original measurements. Also, even if no visually identifiable diagnostic clues appeared, it might be of interest to analyze the statistics of the ICs in different phases of the exercise test with respect to the known heart conditions. However, possible diagnostic value of the proposed method is not evaluated in this paper.

By an appropriate parameterization, it is possible to process the whole section of interest of the ECG exercise test with one run of the ICA algorithm, like was done in our examples. Furthermore, parameterized ECG signals and their ICs for the entire test can be displayed in one screenfull, while traditionally a cardiologist browses through lengthy ECG printouts in search of diagnostic clues. The parameterized ECG signals seen in Figs. 2A and B correspond to roughly 22.5 m of actual ECG on paper.

The proposed method may be of great use in a vast number of applications, for example, in automated or human operated diagnostic and identification systems, for which a meaningful parameterization of the original signals can be found, while retaining the ICA signal model.

References

[1] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York, NY, USA: John Wiley & Sons, 2001.

[2] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Networks*, vol. 10, pp. 626–634, May 1999.

[3] The FastICA package for Matlab.Helsinki University of Technology, Espoo, Finland, 2005. Available: http://www.cis.hut.fi/projects/ica/fastica/.

[4] V. Zarzoso and A. K. Nandi, "Noninvasive fetal electrocardiogram extraction: blind separation versus adaptive noise cancellation," *IEEE Trans. Biomed. Eng.*, vol. 48, pp. 12–18, Jan. 2001.

[5] F. Castells, J. J. Rieta, J. Millet, and V. Zarzoso, "Spatiotemporal blind source separation approach to atrial activity estimation in atrial tachyarrhythmias," *IEEE Trans. Biomed. Eng.*, vol. 52, pp. 258–267, Feb. 2005.

[6] L. Sörnmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. London, UK: Elsevier Academic Press, 2005.

[7] R. Lehtinen, J. Sievänen, V. Turjanmaa, K. Niemelä, and J. Malmivuo, "Effect of ST segment measurement point on performance of exercise ECG analysis," *Int. J. Cardiol.*, vol. 61, pp. 239–245, Oct. 1997.

[8] J. Viik, R. Lehtinen, V. Turjanmaa, K. Niemelä, and J. Malmivuo, "Correct utilization of exercise electrocardiographic leads in differentiation of men with coronary artery disease from patients with a low likelihood of coronary artery disease using peak exercise ST-segment depression," *Am. J. Cardiol.*, vol. 81, pp. 964–969, Apr. 1998.

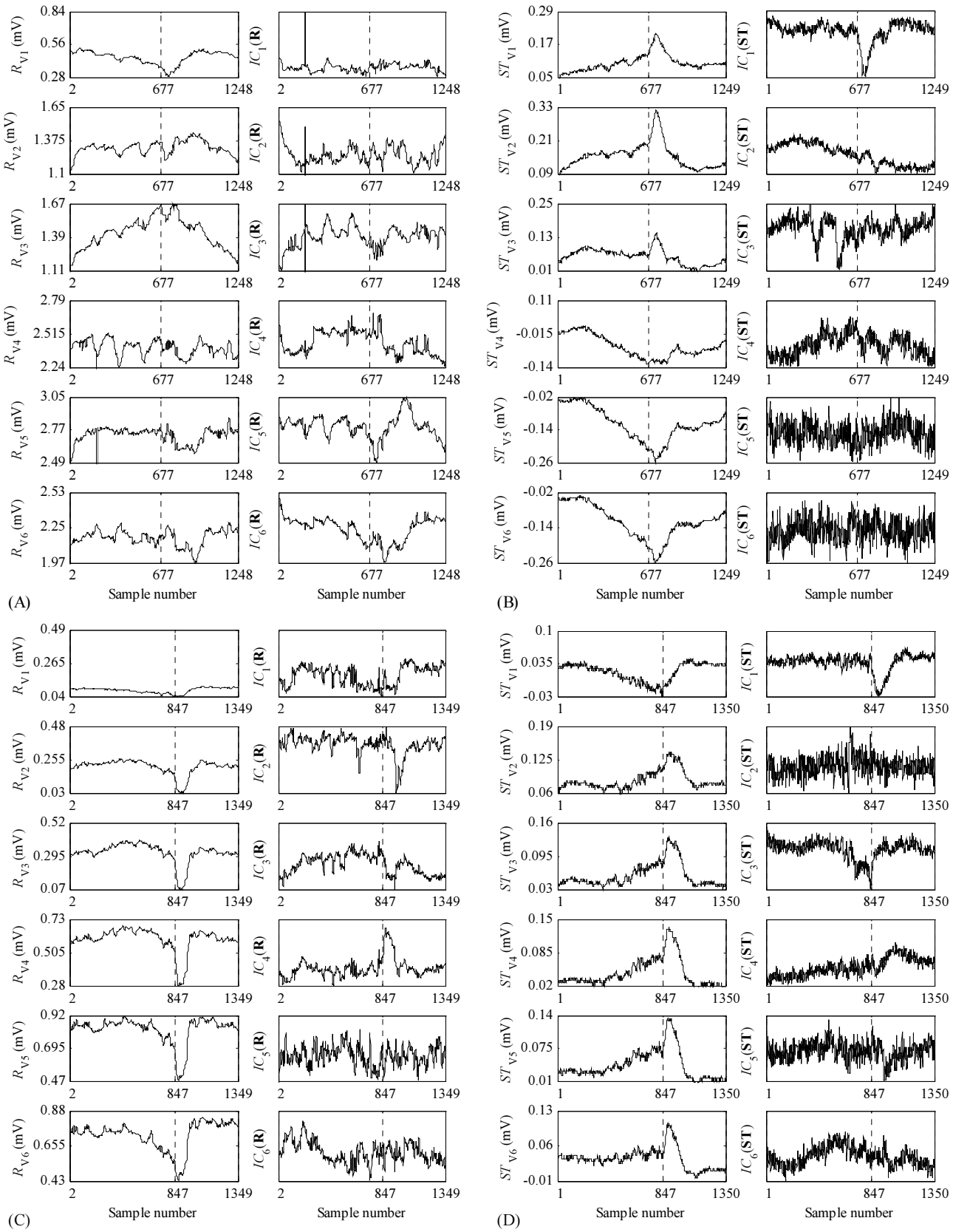

Fig. 2. (A) In the right column, R parameterizations of the ECG signals seen in part in Fig. 1A for the exercise and recovery phases (approximately 15 min altogether) are shown. The corresponding ICs are shown in the right column. (B) The ST parameterizations and the corresponding ICs from the same period and patient as in (A). (C) and (D) The R and ST parameterizations from another patient, respectively, and the corresponding ICs. The end of the exercise phase is indicated by the dashed vertical line. The exercise phase followed by the recovery phase. Note that the amplitude range shown in each subfigure for a parameter signal within the set of six, is always the same, although the scales are different. The amplitudes and signs of the ICs, and the IC order, are arbitrary. Sample number corresponds to the number of the heart beat complex, sample number one being the first collected during the exercise phase.