Comparison of JADE and Canonical Correlation Analysis for ECG de-noising

Jakub Kuzilek[1], Vaclav Kremen[1,2] and Lenka Lhotska[1]

***Abstract*— This paper explores differences between two methods for Blind Source Separation within frame of ECG de-noising. First method is Joint Approximate Diagonalization of Eigenmatrices, which is based on estimation of fourth order cross-cummulant tensor and its diagonalization. Second one is the statistical method known as Canonical Correlation Analysis, which is based on estimation of correlation matrices between two multidimensional variables. Both methods were used within method, which combines the Blind Source Separation algorithm with decision tree. The evaluation was made on large database of 382 long-term ECG signals and the results were examined. Biggest difference was found in results of 50 Hz power line interference where the CCA algorithm completely failed. Thus main power of CCA lies in estimation of unstructured noise within ECG. JADE algorithm has larger computational complexity thus the CCA perfomend faster when estimating the components.**

I. INTRODUCTION

In 1998 *Wisbec et al.* published a manuscript describing deployment of ICA on ECG obtained from 8 precordial electrodes [1]. In the same year *Barros et al.*[2] presented their contribution on ECG source separation using ICA neural network implementation. Following these two pilot works in discussed field other researchers provided their solutions [3], [4], [5], [6], [7], [8].

There exist several works from area of functional magnetic resonance imaging (fMRI). *Thomas et. al* [9] proposed a solution for noise reduction of noise within BOLD-based fMRI using Principal and Independent Component Analysis (PCA and ICA). Another study related to our work was reported by *Liu et al.* [10]the researchers used Canonical Correlation Analysis (CCA) with Singular Value Decomposition (SVD) to reduce noise contained in fMRI.

This paper deals with ECG de-noising using two different methods - Join Approximate Diagonalization of Eigen matrices (JADE) and Canonical Correlation Analysis (CCA). Both methods were applied on data using methodology described in [11]. This paper explores main differences between these two methods, when applied on ECG data with low number of recorded leads (holter ECG, etc.).

II. DATA

Evaluation database was formed by signals from databases available on MIT data storage Physionet [12]: Normal Sinus Rhythm database, European ST-T database, Long Term ST database, QT database, MIT Long Term database and MIT-BIH ST Change database. This gives us a database containing 382 ECG recordings from different sources. All recordings were resampled to sampling frequency 500 Hz in order to make the evaluation easy to interpret.

III. METHODS

A. Blind Source Separation

For "detection" of independent sources in our case - ECG and noise presented in ECG recordings one can use several different approaches. Each is solution of the Blind Source Separation problem (BSS), which can be defined as extraction of a signal set based on knowledge of their mixtures only[13]. Basic BSS model assumes linear combination mixing signals (components):

$$\mathbf{X} = \mathbf{AS}, \tag{1}$$

where $\mathbf{X}$ is a mixture of source signals $\mathbf{S}$ mixed by matrix $\mathbf{A}$. Mixture matrix $\mathbf{A}$ is considered squared, which does not need to be in general. Components are then obtained:

$$\mathbf{S} = \mathbf{A}^{-1}\mathbf{X} = \mathbf{WX}, \tag{2}$$

where matrix $\mathbf{W}$ is inverse to matrix $\mathbf{A}$. Estimation of a components can be reduced to search of matrix $\mathbf{W}$. In general all BSS methods estimate source signals that are as independent/uncorrelated as possible.

B. Joint Approximate Diagonalization of Eigen matrices

Joint Approximate Diagonalization of Eigen matrices algorithm (*JADE*) is blind source separation method based on diagonalization of fourth-order cumulant tensor. This tensor contains all fourth-order information [13]. The cumulant tensor is a linear operator defined by the fourth-order cumulants defining linear transformation in the space of $n \times n$ matrices. The i, jth element of matrix given by transformation $\mathbf{F}_{ij}$ is defined as [13]:$\mathbf{F}_{ij}(\mathbf{M}) = \sum_{kl} m_{kl} cum(x_i, x_j, x_k, x_l)$, where m_{kl} are the elements of transformed matrix $\mathbf{M}$. Considering whitened data: $\mathbf{z} = \mathbf{Vx} = \mathbf{VAs} = \mathbf{W}^T\mathbf{s}$, where $\mathbf{V}$ is whitening matrix, $\mathbf{x}$ are mixed signals, $\mathbf{A}$ is mixing matrix and $\mathbf{s}$ are source signals, the cumulant tensor of $\mathbf{z}$ has structure that can be interpreted as the eigenvalue decomposition. It can be shown that [13]:

$$\mathbf{F}_{ij}(\mathbf{w}_m\mathbf{w}_m^T) = w_{mi}w_{mj}kurt(s_m), \tag{3}$$

where $kurt(s_m)$ is kurtosis of corresponding source signal s_m. If we knew eigenmatrices we could easily obtain independent components. In real world problems the model does

[1]Jakub Kuzilek, Vaclav Kremen and Lenka Lhotska are with the Department of Cybernetics, FEE, Czech Technical University in Prague, Prague, 16627, contact e-mail: jakub.kuzilek at fel.cvut.cz

[2]Vaclav Kremen is with Czech Institute of Informatics, Robotics, and Cybernetics, CTU in Prague, Prague, 166 27

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

not hold exactly and exact diagonalization is almost impossible [13]. Thus during diagonalization measuring diagonality of matrix $\mathbf{WF}(\mathbf{M})_i\mathbf{W}^T$ is needed[13]:

$$J_{JADE}(\mathbf{W}) = \sum_i ||diag(\mathbf{WF}(\mathbf{M})_i\mathbf{W}^T)||^2, \tag{4}$$

where $||diag()||^2$ means sum of squares of the diagonal.
The crucial is selection of set of matrices $\mathbf{M}_i$. One possible choice is to take the eigenmatrices of the cumulant tensor, which is exactly what JADE algorithm does [13].

C. Canonical Correlation Analysis

Canonical Correlation Analysis (CCA) is statistical technique developed by H. Hotelling [cit 10], which measures linear relationship between two multidimensional variables. CCA is invariant to affine transformations of the variables. Let assume measured signals $x_i(t)$ and their time delayed versions $x_i(t-k)$ now we can define k-step autocorrelation of component $s_i(t) = W_iX(t)$, where W_i is corresponding row of de-mixing matrix, can be expressed as follows:

$$\rho_i = \frac{W_iC_{x(t)x(t-k)}W_i}{\sqrt{W_iC_{x(t)x(t)}W_i}\sqrt{W_iC_{x(t-k)x(t-k)}W_i}}, \tag{5}$$

where $C_{x(t)x(t)}$ denotes covariance matrix of vector $x(t)$. $C_{x(t)x(t-k)}$ and $C_{x(t-k)x(t-k)}$ is defined in similar way. Autocorrelation reaches maximum when W_i satisfies equation:

$$C^{-1}_{x(t)x(t)}C_{x(t)x(t-k)}C^{-1}_{x(t-k)x(t-k)}C_{x(t-k)x(t)}W_i = \rho_i^2W_i. \tag{6}$$

We can observe that ρ_i^2 is an eigenvalue of of matrix $C^{-1}_{x(t)x(t)}C_{x(t)x(t-k)}C^{-1}_{x(t-k)x(t-k)}C_{x(t-k)x(t)}$ and W_i is an eigenvector associated with corresponding eigenvalue, thus we can find several pairs of vectors $[W_1, ..., W_L]$, where L is number of estimated mutually uncorrelated components s_i. Covariance matrices can be estimated as:

$$C_{xy} = \frac{1}{T}\sum_{t=1}^{T} x(t)y(t), \tag{7}$$

where x and y are corresponding vectors - in our case $x(t)$ or $x(t-k)$. For our purposes we used delay $k = 1$.

D. BSS for de-noising

For testing we used framework proposed in [11] the ECG de-noising algorithm combines BSS with CART decision tree in order to identify and remove noise from ECG. The algorithm work-flow is shown in Figure 1. Firstly mean is subtracted from ECG recording. Then the independent components using BSS (JADE or CCA) are estimated. For each component set of features were computed:

- Mean of component
- Variance of component
- Kurtosis of component
- Standard deviation of component peak-to-peak distances

These features are passed to trained decision tree (for training Gini's impurity criterion and cross-validation pruning were used), which decides whether component is noise or ECG

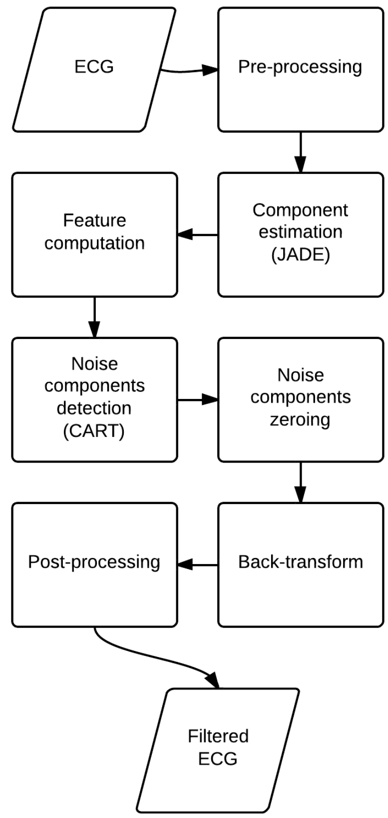

Fig. 1. ICA based noise removal algorithm work-flow. First data is preprocessed, then the components are estimated. After that components containing noise are identified and removed. Finally ECG is reconstructed and filtered using post-processing filter.

containing. The tree is trained using components obtained using MIT/BIH Arrhythmia database and artificially generated noise. All training components were annotated and then used for training the decision tree.
Finally components marked as noise are removed, all components are projected back to signal domain, filtered to remove high frequency noise (observed on 4 cases of training database) using low pass filter with first zero at 117 Hz, delay 5 samples and gain 0.93.

E. Evaluation

For evaluation we used Root-Mean-Square Error ($RMSE$), which is good statistical index for case, when the original clear signal is known. $RMSE$ is defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}(x_i - y_i)^2}, \tag{8}$$

where x_i is the i-th sample of original signal, y_i is the i-th sample of filtered signal and N is the number of samples in both signals. $RMSE$ equal to zero means that original and filtered signal are identical.
During the evaluation process several steps were passed in order to explore performance of methods. First the data

were mixed with the simulated noise: EMG noise (random Gaussian signal), Power line interference (50 Hz sinusoid), Baseline wander (0.333 Hz sinusoid), Electrode cable movement (sum of sinusoids with amplitudes and frequencies ranging from 0.1 to 1 mV and 1.5 to 8 Hz). Each type of noise is added to ECG at 25, 50, 75, 100 percent of the maximum noise amplitude levels.

Then de-noising was applied on the data and the results were obtained. Finally resulting signals were compared to the original ECG data using $RMSE$ measure and the result is stored for the evaluation.

IV. RESULTS AND DISCUSSION

We performed statistical testing of $RMSE$ obtained as a result of comparison between original and filtered signals. The results are summarized using boxplot figures (Fig. 2, 3, 4, 5). Each figure contains 10 boxplots, which are linked to one type of noise. Boxplots are grouped according to the level of noise added to the original recording during the testing. Each boxplot has "notches", which shows 95% confidence intervals for null hypothesis that $RMSE$ differs from others. If the confidence intervals are overlapping then we reject the null hypothesis.

Figure 2 shows the comparison of results on ECG contaminated by electrode cable artefact. We can observe that JADE based de-noising performed better than CCA based de-noising. Specially in case of noise with higher amplitude. The main reason for this is that electrode cable movement artefact is structured type of noise and CCA is good for uncovering structured elements of signals. Thus it has problems when noise is structured, because ECG is structured too.

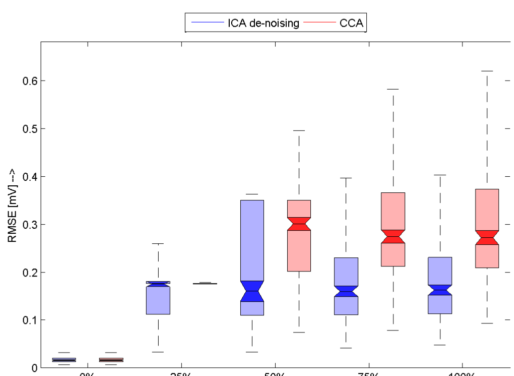

Fig. 2. Results for the electrode cable movement artefact. Horizontal axis shows level of added noise. Boxplots shows RMSE for used database of signals.

Figure 3 shows the results for ECG mixed with EMG. We can observe that both algorithms performed in similar way. The EMG noise was simulated as white noise thus it has no structure and the ECG containing components were identified and noise is reduced in both cases.

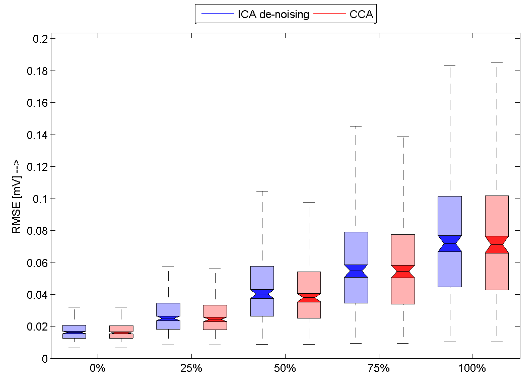

Fig. 3. Results for EMG artefact. Horizontal axis shows level of added noise. Boxplots shows RMSE for used database of signals.

Figure 4 shows the results of ECG modulated by base line wander. Again we can see that JADE is slightly better than CCA, but the performance of both methods is in general the same.

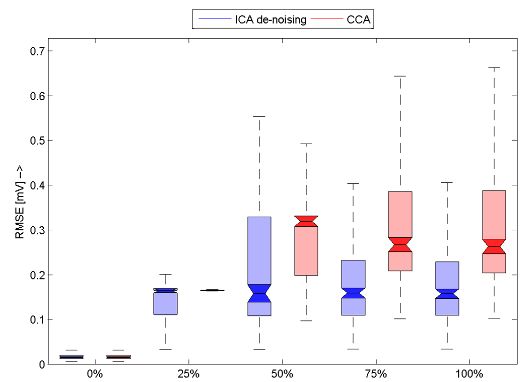

Fig. 4. Results for the base line wander artefact. Horizontal axis shows level of added noise. Boxplots shows RMSE for used database of signals.

Finally figure 5 shows the results on recordings contaminated by power line interference. The CCA is failing in this case in reduction of this type of noise. This is most probably due to simulation of noise using 50 Hz sinus, but this also shows the biggest problem of method in application to ECG de-noising - structured noise cannot be efficiently distinguished from the ECG activity.

Figures 6 and 7 show the main difference between JADE component estimation and CCA component estimation. Because the CCA uses the correlation as a measure of "independence" it is unable to separate power line interference and ECG - both of them hasse large correlation and their mixture has it too, thus CCA, which is based on uncovering structures in signals, cannot efficiently separate ECG and noise (Fig. 7). On the other hand ECG has super-Gaussian distribution [3] thus kurtosis based JADE is able to separate

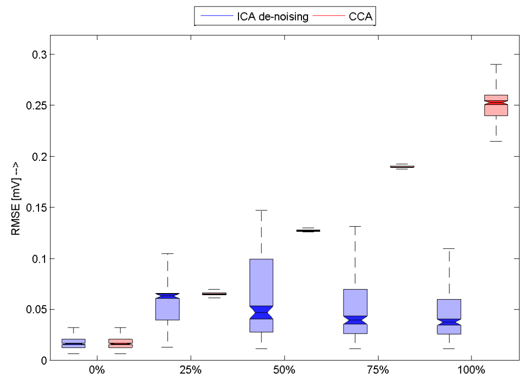

Fig. 5. Results for the power line interference artefact. Horizontal axis shows level of added noise. Boxplots shows RMSE for used database of signals.

ECG and noise to different components (Fig. 6).

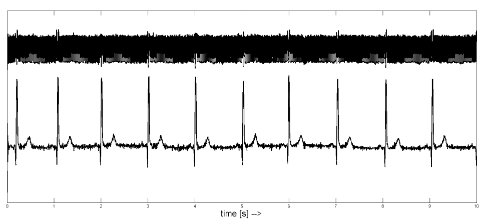

Fig. 6. Example of estimated ECG components using JADE algorithm.

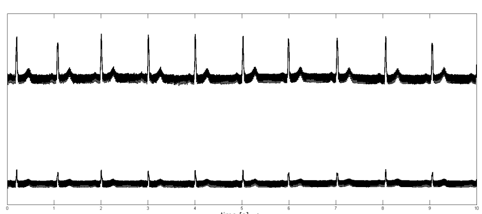

Fig. 7. Example of estimated ECG components using Canonical Correlation Analysis.

V. CONCLUSIONS

This work uncovers differences between CCA and JADE, we observed that in case of unstructured noise (EMG) CCA performance was similar to JADE. On the other hand when noise shows any structure the CCA starts to have difficulties in separation of noise and ECG. We observed that biggest difference in results was in case of 50 Hz power line interference where the CCA algorithm completely failed. Thus main power of CCA lies in estimation of unstructured noise within ECG. The second difference, which has been observed during the experiments that the JADE algorithm has larger computational complexity due the estimation of kurtosis and Given's rotations used for estimation of independent components. Thus the CCA is performing faster when estimating the components.

BSS methods suffer in general with environment changes and thus more precise testing needs to be implemented. We can assume that the general structure of the algorithm remains unchanged because the main idea is to search for the ECG activity within noise and thus we can assume that the ECG characteristics will remain the same. We are also planning experiments with segmentation and its effect on algorithm performance.

ACKNOWLEDGMENT

This work was supported by post doctoral research project by Czech Science Foundation GACR #P103/11/P106 and research program No. MSM 6840770012 "Transdisciplinary Research in Biomedical Engineering II" of the CTU in Prague.

REFERENCES

[1] J. O. Wisbeck, A. K. Barros, A. K. B. Yy, and R. G. Ojeda, "Application of ica in the separation of breathing artifacts in ecg signals," 1998.

[2] A. K. Barros, A. Mansour, and N. Ohnishi, "Removing artifacts from electrocardiographic signals using independent components analysis," *Neurocomputing*, vol. 22, no. 1-3, pp. 173–186, 11/20 1998.

[3] T. He, G. Clifford, and L. Tarassenko, "Application of independent component analysis in removing artefacts from the electrocardiogram," *Neural Computing & Applications*, vol. 15, no. 2, pp. 105–116, 2006.

[4] M. P. S. Chawla, H. K. Verma, and V. Kumar, "Artifacts and noise removal in electrocardiograms using independent component analysis," *International journal of cardiology*, vol. 129, no. 2, pp. 278–281, 9/26 2008.

[5] M. Milanesi, N. Martini, N. Vanello, V. Positano, M. F. Santarelli, and L. Landini, "Independent component analysis applied to the removal of motion artifacts from electrocardiographic signals," *Medical & biological engineering & computing*, vol. 46, no. 3, pp. 251–261, MAR 2008.

[6] D. DiPietroPaolo, H.-P. Mueller, G. Nolte, and S. N. Erne, "Noise reduction in magnetocardiography by singular value decomposition and independent component analysis," *MEDICAL & BIOLOGICAL ENGINEERING & COMPUTING*, vol. 44, no. 6, pp. 489–499, JUN 2006.

[7] Y. Tu, X. Fu, D. Li, C. Huang, Y. Tang, S. Ye, and H. Chen, "A Novel Method for Automatic Identification of Motion Artifact Beats in ECG Recordings," *ANNALS OF BIOMEDICAL ENGINEERING*, vol. 40, no. 9, pp. 1917–1928, SEP 2012.

[8] J. Oster, O. Pietquin, R. Abaecherli, M. Kraemer, and J. Felblinger, "Independent component analysis-based artefact reduction: application to the electrocardiogram for improved magnetic resonance imaging triggering," *PHYSIOLOGICAL MEASUREMENT*, vol. 30, no. 12, pp. 1381–1397, DEC 2009.

[9] C. G. Thomas, R. A. Harshman, and R. S. Menon, "Noise reduction in bold-based fmri using component analysis," *NeuroImage*, vol. 17, no. 3, pp. 1521 – 1537, 2002.

[10] Y. Liu, D. Hu, Z. Zhou, H. Shen, and X. Wang, "fmri noise reduction based on canonical correlation analysis and surrogate test," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, no. 6, pp. 870–878, Dec 2008.

[11] J. Kuzilek, "Independent component analysis: Applications in ecg signal processing," Ph.D. dissertation, Czech Technical University, 2013.

[12] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet : Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. e215–220, Jun. 2000.

[13] A. Hyvrinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Wiley, 2001.