1. Tomasz RYMARCZYK^{1,2}, 2. Andres VEJAR^{1,2}, 3. Timur KASSYMOV¹

Research & Development Center Netrix S.A., 20-704 Lublin, Poland (1), Institute of Computer Science and Innovative Technologies, University of Economics and Innovation in Lublin, 20-209 Lublin, Poland (2) ORCID: 1. 0000-0002-3524-9151; 2. 0000-0002-2041-0387, 3. 0000-0002-1171-7467

doi:10.15199/48.2022.03.25

Concurrent data acquisition of electrocardiogram and respiration rate using predictive filtering

Abstract. For e-health monitoring, concurrent data acquisition and processing of electrocardiogram and respiratory rate biosignals is needed to provide clean input signals for the detection algorithms. Given that cardiac and respiration rhythms are correlated, but they have different characteristic rates, we explored a sensor fusion approach to model the filtered future state of both signals. Therefore, a time-delayed structure with raw and filtered data fusion was used to improve the quality of the model. We apply Lasso and Ridge regression to the time-delayed data structure and study the approach using low and high sample rates for the data processing. At low sample rates, the validation for respiratory rate is high, while for electrocardiogram is comparatively low. With higher sample rates, the results are improved for both cases.

Streszczenie. W przypadku monitorowania e-zdrowia, jednoczesne pozyskiwanie i przetwarzanie danych z elektrokardiogramu i biosygnałów oddechowych jest konieczne, aby zapewnić czyste sygnały wejściowe dla algorytmów detekcji. Biorąc pod uwagę, że rytmy serca i oddychania są skorelowane, ale mają różne charakterystyczne prędkości, zbadaliśmy podejście do fuzji czujników, aby modelować przefiltrowany przyszły stan obu sygnałów. Dlatego też, w celu poprawy jakości modelu, zastosowano strukturę czasowo-opóźnioną z fuzją danych surowych i przefiltrowanych. Zastosowaliśmy regresję Lasso i Ridge do struktury danych opóźnionych w czasie i zbadaliśmy podejście do duzji czytowania, walidacja dla częstości oddechów jest wysoka, podczas gdy dla elektrokardiogramu jest porównywalnie niska. Przy wyższych częstotliwościach próbkowania, wyniki są lepsze dla obu przypadków. (Jednoczesna akwizycja danych elektrokardiogramu i częstości oddechów z wykorzystaniem filtrowania, mejstewania, wej stanie mista.

Keywords: predictive filtering, biosignals, sensors. Słowa kluczowe: filtrowanie predykcyjne, biosygnały, czujniki.

Introduction

Modern cardiorespiratory monitoring and detection can be used to support standard diagnostic methods to improve the allocation of scarce resources and personnel in the medical environment. Current respiratory examination methods can be enhanced in many aspects, especially data accuracy [1]. The main problem is unexpected dynamics, caused internally or externally. As reviewed in [2], long-term monitoring is difficult and costly, and automation is required. Considering online processing of medical biosignals [3], predictive filtering is an interesting approach. This work is the continuation of [4], where a joint linear regression using the Ridge and Lasso method for one-step-ahead prediction of electrocardiogram (ECG) and respiration rate (RR) was analyzed for concurrent data acquisition of these signals. An example of ECG and RR is presented in Figure 1.



Fig.1. Sample of concurrent acquisition of raw ECG and raw RR (top) biosignals and detail on filtered signals.

This project aims to predict the future state of both signals and estimate the filter output in the current instant for cardio-respiratory monitoring. In Figure 2, we present the application concept.

Sensor Data Fusion

Multi-sensor data fusion can be defined as a complex methodology, techniques, and tools used to combine sensor data into a common representation. The main objective of sensor fusion practice is information quality improvement.



Fig.2. ECG and RR data fusion for cardio-respiratory health monitoring.

A research method, multi-sensory data fusion, has become important. It is a working process of operating on multiple source data and information combinations that is required to improve the accuracy and extend the inferences concerning the use of a single, specific sensor. However, this cannot be considered a novelty. Examples of multiple sensors usage can be observed in living systems. As evidence, it may not be possible to assess the quality of an edible substance based solely on the sense of vision or touch, but evaluation of edibility may be achieved using a combination of sight, touch, smell, and taste. Furthermore, while one is unable to see around comers or through vegetation, the sense of hearing can provide advanced warning of impending dangers. Therefore, the fusion of data is naturally performed by living organisms to achieve a more accurate assessment of the environment and identification of threats, improving the chance of survival [5].

Data fusion can improve the performance of the system in four different ways:

- 1. Representation. The new granularity provides data with a richer than each of the initial sources of information.
- 2. Certainty. If *D* is a sensor data before fusion and $p\{D\}$ is the a priori probability of the data before fusion, then the gain in certainty is the growth in $p\{D\}$ after fusion. If D_F denotes data after fusion, then we expect $p\{D_F\} > p\{D\}$.
- 3. Accuracy. The standard deviation on the data after the fusion process is smaller than the standard deviation provided directly by the sources. Thus, if data is noisy or erroneous, the fusion process tries to reduce or eliminate it.
- 4. Completeness. If the information is redundant and concordant, we could also have a gain inaccuracy [6].

Different sensors and measurement techniques can be added up in cardiorespiratory monitoring using fusion. Advanced multielectrode methods can be considered, as in electrical tomography [7, 8, 9]. Also, advanced indicators can be studied, e.g., mutual information [10]. In [11], it was reviewed that there is a lack of a single-lead ECG technique that would be equally applicable for all population groups: the respiratory sinus arrhythmia is often less prominent among seniors, and R-peak amplitude modulation has high motion sensitiveness. In this range, data fusion has previously been applied to the estimation of respiratory rate to compensate for noisy measurements when multiple physiological signals are a-vailable (such as the photoplethysmograph, impedance pneumograph, arterial blood pressure and peripheral arterial tonometry waveform). Note that this requires at least two different signals to be available. The main idea is that considering respirationinduced ECG modulation is generated by two independent processes (autonomic control and chest movement), with different modes of failure, integrating information from both sources should result in an improved respiratory rate estimate. Data fusion can be applied for other medical purposes. For example, [12] presented a novel method for seizure onset detection. It uses using fused information extracted from multichannel electroencephalogram and single-channel electrocardiogram.

Cardio-respiratory Health Monitoring

This study aims to predict the filtered future state of both signals using a data fusion method for cardiopulmonary monitoring. Figure 2 shows the data fusion concept for this application.

The raw patient signals are measured in the ECG Data acquisition from surface potentiometers and RR from the accelerometer. The *Joint analysis* block then considers a chain of modules:

- 1. *Predictive filter* We model the filtered future state of both signals using a time delay structure. Raw and filtered ECG and RR sensor data were used to verify the quality of the model. We use LASSO and Ridge regression to estimate the model parameters.
- 2. The *Signal averaging* module is a filter that converts the measured signal to the desired ratio for the detection or diagnostic module.
- 3. The *Detection diagnosis* module evaluates the signals to diagnose a disease or detect abnormal cardiopulmonary behaviour.

Finally, in the *Estimation* block, the alarm system can take into account the severity level set by the *Detection diagnosis* module and send an alarm to the corresponding monitoring system.

Predictive Filtering

Predictive filtering is a type of adaptive filtering that is robust to dynamic changes in the patient-sensor interface. For example, spontaneous activity and body motion usually cause the filter parameters to change adaptively based on online signals, so the result is stable under a certain level of exogenous noise. Adaptive filters also have a wide range of applications [13]. In Figure 3, we show the application concept in a condensed form.

ECG and RR raw signals are measured concurrently from the patient. We designate the ECG as $x_1(t) \in \mathbb{R}$ and RR as $x_2(t) \in \mathbb{R}$. The measured signals are subject to the disturbances n_1 and n_2 , and are specified by $x_{1m} = x_1 + n_1$ and $x_{2m} = x_2 + n_2$.



Fig.3. Predictive filter with two delays, block diagram. The ECG and the RR biosignals are associated with the variables x1 and x2, respectively.

For the diagram of Figure 3, the continuous time, continuous valued variable $x(t) = [x_{1m}(t), x_{2m}(t)]^T$, is sampled with sampling period T_s in its discrete timed version $x[k] = x(T_sk) \in \mathbb{R}^2$. The predictive filter G_{θ} is defined by:

(1) $y[k] := x_{F}[k] = G_{\theta}(x[k-1], x[k-2], x_{F}[k-1], x_{F}[k-2])$

The delayed filtered values are calculated by a linear filter H that is used to calculate the delayed version of the filtered signals:

(2)
$$x_F[k] = H(x[k], x[k-1])$$

The principle of linear frequency filtering is based on the difference in the spectral characteristics of biosignals and various types of interference. The use of bandpass frequency filtering at the biomedical signal preprocessing stage allows reducing the influence of various kinds of interference, both of physical and biological origin [14].

The parameter update of the filter, specified as θ , represents the transformation of the input data (x[k-1],x[k-2]) into the output data y, given the defined structure of the function G_{θ} . In the linear case, without considering the delayed filtered signals, we can select a prediction structure for $G_{\theta} : \mathbb{R}^{2 \times} \mathbb{R}^{2} \to \mathbb{R}^{2}$ as $\theta_{0} + \theta_{1}x[k-1] + \theta_{2}x[k-2]$.

The sum of square errors ${}^{P}_{i} e^{2}_{i} = {}^{P}_{i} |x_{i}-y_{i}|^{2}$ is generally used to obtain a θ^{2} that reduces the error given a subsample of data collected. This can be done iteratively, estimating θ^{2} at every timestep as in the recursive least squares filter, or updating θ^{2} every *W* timesteps.

The current state of the signals is taken into account to obtain a prediction of filtered signals. The result is an evaluation of the signals for diagnosing any disease or detecting abnormal cardiovascular behaviour, as shown in Figure 2.

Cardiorespiratory Data Acquisition

Electrocardiography can conventionally be described as a recording of the propagation of the electric field through the tissue structures of the heart, correcting for the dimensionality of this process. Electrodes are placed on the surface of the patient's body and detect small changes in skin potential due to depolarization of the heart muscle with each contraction. Portable devices for monitoring the health of the heart can thus be created on the basis of the electrical modulation responses of the heart: ECGs, cardiac monitors, etc. In addition, the same principle is used to obtain data on contractions of other muscles, as is the case with the electromyogram. Electrical potentials can be used in bionics and prosthetics. The electrodes must be positioned so that they are directed towards the muscle whose activity is to be monitored. In this project, we will use one of the standard ECG lead diagrams shown in Figure 4.



Fig.4. Left: Two common 3-leads ECG electrode configurations. Right: accelerometer location to track respiration activity.

To collect ECG and RR data simultaneously, a combined design was made as shown in Figure 5. Both the ECG and RR sensor are connected to the microcontroller, and the whole device is connected to the client computer via a universal serial bus. For data acquisition, we use the ESP8266 microcontroller from Espressif Systems. The ECG is performed with an AD8232 device from Analog Devices. For the RR, the LSM9DS1 inertial measurement unit from STMicroelectronics was chosen.



Fig.5. Measuring device (left). Measuring system environment (right).

From the right side of Figure 5, it is possible to estimate the position of the accelerometer. It was decided to place the accelerometer at the point between the chest and the abdomen since the instrument can be kept stable in this position and obtain the correct readings. The greatest displacement during breathing occurs on the x-axis of the accelerometer. All breaths, all breathing cycles, their amplitude, and the time they were taken and are visible on this axis. The y and z axes can be used to detect user activities other than breathing. At this point, the activity sensor is mounted on the bottom edge of the circuit board, making it suitable for measuring iris activity.

The implementation of the work consists of ten following steps:

- 1. Data acquisition.
- 2. Transferring data to the PC.
- 3. Setting delays and specific data segments.
- 4. Machine learning.
- 5. Making a threshold model with a Ridge/Lasso linear regression.

- 6. Training, validating and plotting.
- 7. Transferring parameters of predictive filter to the microcontroller.
- 8. Signal filtering update.

Predictive Filter Training with Ridge and LASSO Regression

Various methods can be used to solve the optimisation process [15-21]. We use linear regression to train the predictive filter model and update the predictive parameters online. In particular, we use the Ridge and LASSO regression methods. for prediction [22]. The most well-known machine learning methods are ridge regression and LASSO. Ridge and LASSO are among the most well-known regression methods for machine learning. They can be used to reduce model complexity and avoid the overlearning that can occur with simple linear regression. If we represent the problem as a linear model based on p number of functions, the information will be as signed a hyperplane over the explained variable y with intercept b:

$$(3) y = a_1x_1 + \cdots + a_px_p + b$$

A sample S of N data points of the hyperplane is presented in the form:

(4)
$$\mathcal{S} = \{(y_i, \{x_{ij}\}_{j=1}^p)\}_{i=1}^N$$

We can associate the data to some observation model, with $\varepsilon \sim N(0,\sigma)$ a stochastic variable that models the data disturbances:

(5)
$$\hat{y}_i = y_i + \varepsilon$$

In order to obtain appropriate estimation parameters a_i from the model (3), it is necessary to reduce the difference between regression calculation y^{*} and the measured data y. This possibility choice is modelled by a cost function f(a), which allows the solution of the hyperplane flags to be explored such that it reduce anomalies in the overall data set S. An almost universal choice is the penalty of the cost function w in the form of an error square with an additional regular expression R(a) to account for certain prior information or design features of solutions a_i .

The differences in the estimates can be seen in Figure 6. The raw data are shown in the background, and the regression results are shown in the foreground. At the left ridge represents a near-perfect fit to the measured values, and on the right is the LASSO-calculated average of these values. After performing the experiments, we chose Ridge regression to update the prediction. Since the same value of the regularization parameter λ , it can predict correctly, whereas the LASSO method averages the values because it sets many parameters to zero due to the LASSO method. L1 norm in control in part. Ridge regression is more robust to real-time estimation in the context of the predictive filtering problem.



Fig.6. Validation through Ridge (left) and LASSO (right) regressions.

In Figure 7, we present the general predictive filter model G:

(6)
$$y_a[k] := x_F^{k}[k + a] = G(\sim x_D, \sim x_S)$$

With *the* number of steps ahead in the prediction, the vectors $\sim xD$ and $\sim xS$ represent the delayed data to buffer, raw and smoothed, to feed the machine learning engine. Thus, the base filter corresponds to a linear filter F_0 .



Fig.7. General diagram of the predictive filter implementation.

In Figure 7, (B_D) stands for raw data buffer, F_0 for linear filter and (B_S) for the online filter. Buffer, F0 - linear filter, (BS) online filter (smoothed) data buffer, (B_{SM}) - smoothed online learning buffer for machine learning and (B_{DM}) - buffer for raw data for machine learning. Machine learning buffer. Block M represents the evaluation of the machine learning model that generates and updates. Parameters θ for estimating the predictive filtering specified in block G. Y_a is the filtered prediction signal. Signal, the output signal of the model. Figure 8 shows the results of the filtered and predicted data for ECG and RR.



Fig.8. ECG and RR predictive filter and filtered original.

Authors: Tomasz Rymarczyk, Prof. Eng., University of Economics and Innovation, Projektowa 4, Lublin, Poland, Research and Development Center Netrix S.A., Wojciechowska 31, Lublin, Poland, E-mail: tomasz.rymarczyk@netrix.com.pl; Ph.D. Eng. Andres Vejar, Netrix S.A. Research and Development Center, ul. Związkowa 26, 20–148 Lublin, Poland, University of Economics and Innovation in Lublin, ul. Projektowa 4, 20–209 Lublin, Poland. E-mail: andres.vejar@netrix.com.pl. Eng. Timur Kassymov, Netrix S.A. Research and Development Center, ul. Związkowa 26, 20– 148 Lublin, Poland.

REFERENCES

[1] Philip K. E., Pack E., Cambiano V., Rollmann H., Weil S. and O'Beirne J., The accuracy of respiratory rate assessment by doctors in a london teaching hospital: a cross-sectional study, *Journal of clinical monitoring and computing*, 29 (2015), No. 4, 455–460

- [2] Serhani M. A., Kassabi H. T. El, Ismail H., Navaz A. N., Ecg monitoring systems: Review, architecture, processes, and key challenges, *Sensors*, 20 (2020), No. 6, 1796
- [3] Bansal D., Khan M., and Salhan A. K., A computer based wireless system for online acquisition, monitoring and digital processing of ecg waveforms, *Computers in biology and medicine*, 39 (2009), No. 4, 361–367
- Kassymov T., Concurrent data acquisition and computer analysis of electrocardiogram and respiratory rate signals, March 2021
- [5] Hall D. L. and Llinas J., An introduction to multisensor data fusion, *Proceedings of the IEEE*, 85 (1997), No. 1, 6–23
- [6] Mitchell H. B., Multi-sensor data fusion: an introduction. Springer Science & Business Media, 2007
- [7] Rymarczyk T., Vejar A., Nita P., and P. Tchórzewski, Advanced tomographic platform for real-time image reconstruction and biomedical signal analysis, in 2018 International Interdisciplinary PhD Workshop (IIPhDW), May 2018, 186–190
- [8] Rymarczyk T., Nita P., Vejar A., Stefaniak B., Sikora J., Electrical tomography system for innovative imaging and signal analysis, *Przegląd Elektrotechniczny*, 95 (2019), No. 6, 135– 138
- [9] Rymarczyk T., Vejar A., Multi frequency electrical tomography with re-configurable excitation waveforms, in 2019 Applications of Electromagnetics in Modern Engineering and Medicine (PTZE), June 2019, 198–202
- [10] Vejar A., Rymarczyk T., Paprzycki P., Mutual information and delay embeddings in polysomnography studies, in 2019 International Interdisciplinary PhD Workshop (IIPhDW), May 2019, 89–94
- [11]Orphanidou C., Fleming S., Shah S. A., and Tarassenko L., Data fusion for estimating respiratory rate from a single-lead ecg, *Biomedical Signal Processing and Control*, 8 (2013), No. 1, 98–105.
- [12] Qaraqe M., Ismail M., Serpedin E., and Zulfi H., Epileptic seizure onset detection based on eeg and ecg data fusion, *Epilepsy & Behavior*, 58 (2016), 48–60
- [13] Correa A. G., Orosco L. L., Diez P., and Leber E. L., Adaptive filtering for epileptic event detection in the eeg, *Journal of Medical and Biological Engineering*, 39 (2019), No. 6, 912–918
- [14] Aboy M., Crespo C., McNames J., Bassale J., and Goldstein B., Automatic detection algorithm for physiologic pressure signal components, in *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society, Engineering in Medicine and Biology*, 1 (2002), 196–197
- [15] Pawlowski, S; Plewako, J; Korzeniewska, E, Influence of Structural Defects on the Resistivity and Current Flow Field in Conductive Thin Layers, *Electronics*, 9 (2020), No.12, 2164
- [16] Rymarczyk T., Kłosowski G., Tchórzewski P., Cieplak T., Kozłowski E.: Area monitoring using the ERT method with multisensor electrodes, *Przegląd Elektrotechniczny*, 95 (2019), No.1, 153-156
- [17] Koulountzios P., Rymarczyk T., Soleimani M., A quantitative ultrasonic travel-time tomography system for investigation of liquid compounds elaborations in industrial processes, *Sensors*, 19 (2019), No.23, 5117
- [18] Kłosowski G., Rymarczyk T., Kania K., Świć A., Cieplak T., Maintenance of industrial reactors based on deep learning driven ultrasound tomography, *Eksploatacja i Niezawodnosc – Maintenance and Reliability*, 22 (2020), No.1, 138–147
- [19] Kłosowski G., Rymarczyk T., Cieplak T., Niderla K., Skowron Ł., Quality Assessment of the Neural Algorithms on the Example of EIT-UST Hybrid Tomography, *Sensors*, 20 (2020), No.11, 3324
- [20] Rymarczyk T.: Characterization of the shape of unknown objects by inverse numerical methods, *Przeglad Elektrotechniczny*, 88 (2012), No.7B, 138-140
- [21] Rymarczyk T., Using electrical impedance tomography to monitoring flood banks, 16th International Symposium on Applied Electromagnetics and Mechanics (ISEM) International journal of applied electromagnetics and mechanics, 45 (2014), 489-494
- [22] Roy S. S., Mittal D., Basu A., and Abraham A., Stock market forecasting using lasso linear regression model, in *AfroEuropean Conference for Industrial Advancement*, Springer, (2015), 371–381