Annals of Biomedical Engineering, Vol. 37, No. 9, September 2009 (© 2009) pp. 1871–1876
DOI: 10.1007/s10439-009-9750-x

Biomedical Applications of Piece-Wise Affine Identification for Hybrid Systems

DIEGO LIBERATI[1,2]

[1]Institute of Electronic, Information and Communication Engineering, Italian National Research Council, Milan, Italy; and [2]Dipartimento di Elettronica e Informazione, Politecnico Piazza Leonardo da Vinci 32, 20133 Milan, Italy

(*Received 2 February 2009; accepted 22 June 2009; published online 1 July 2009*)

Abstract—Modeling switching processes for control purpose takes advantage from the Piece-Wise Affine identification of hybrid dynamical systems briefly recalled in this paper. A couple of applications are addressed, namely to discriminate hormone pulses from background noise, in a physiologically switching process, and to identify sleep apneas, as pathological switching among healthy and potentially risky states. Other potential applications are proposed.

Keywords—Neuroendocrine pulsatility, Sleep apneas, Identification, Inference, Pattern recognition, Clustering, *K-means*.

INTRODUCTION

Nonlinear and/or not stationary processes are in everyday life, especially when dealing with ourselves (like in patho-physiology), with surrounding, with manufacturing: strictly speaking, every real process is almost surely not linear and not stationary, indeed! The traditional easy way to understand their dynamics is, when possible, to compute an approximate process that would be linear around the working equilibrium point, stationary in its surrounding, and almost equivalent to the real one in the range of values of interest. At the price of such quantifiable error, one would gain the possibility do deal with better known linear systems, whose even robust control is at hand via many procedure well described in literature.

In several cases, nonetheless, the dynamics of the process embedded in the black box producing the measured data is almost not at all known in its physical or chemical aspects. Modeling is then not anymore a matter of writing the correct differential equations, even in terms of partial derivatives, but to infer a model directly from the data. A typical approach to such a problem is to try a regression, which is easy if it happens that those particular data belongs to a linear stationary, process (which is not granted nor probable at all), or would also be fair if one would guess the correct nonlinear kernel embedded in the data and the deviation from stationary behavior along its time course, which is not so easy too, especially in quite often practical case of multi-dimensional data produced by a hybrid dynamic and logical process.

On the other side, when data are collected via a complex multivariable sampling, like for micro-arrays in genomics, a pre-processing is also needed in order to select the really salient variables to be included in the identification: the use of minimum description length (MDL) principle[3] may help in providing an answer to such a quest. When inference of only logical static relationship among salient variables is of interest, Adaptive Bayesian Networks (ABN) may be then jointly applied with MDL as illustrated in Bosin *et al.*[5]; alternatively, the Logical Networks (LN) proposed in Muselli and Liberati[13] may be directly used.

When proper salient variables are selected, either by means of the recalled techniques, of via prior knowledge on the investigated process, the Piece-Wise Affine (PWA) approach proposed in Ferrari-Trecate *et al.*,[8] and briefly recalled in its core aspects in the following, is a possible easy solution to perform the identification of a dynamic model from data, as it will be discussed in this paper, especially in the study of the human patho-physiology, with reference to a few paradigmatic examples. Such a general approach is quite a powerful tool in order to identify from data not only the switches and the approximated linear dynamics between them in case of intrinsically hybrid dynamic-logic processes, like the one described in this paper, but, not surprisingly, also piece-wise linear approximations (and even hysteretic behavior), as it will also be discussed in a few examples, where other approaches will be recalled in order to easy comparison.

Address correspondence to Diego Liberati, Dipartimento di Elettronica e Informazione, Politecnico Piazza Leonardo da Vinci 32, 20133 Milan, Italy. Electronic mail: liberati@elet.polimi.it

A couple of applications are discussed in the study of the human patho-physiology, namely: identifying pituitary hormone pulses in background noise from sampled time course of blood concentration; detecting apneas from ECG monitoring. Other promising directions of investigation are also addressed.

IDENTIFICATION OF PIECE-WISE AFFINE SYSTEMS THROUGH A CLUSTERING TECHNIQUE

The dynamics of a multivariable systems may be described in a general way as a set of s Piece-Wise linear AutoRegressive with eXogenous input (PWARX) models:

$$y(k) = \begin{cases} \theta'[\mathbf{x}'(k)\ 1]' & \text{if} \quad \mathbf{x}(k) \in X_1 \\ \vdots & \\ \theta_s'[\mathbf{x}'(k)\ 1]' & \text{if} \quad \mathbf{x}(k) \in X_s \end{cases} \tag{1}$$

each one describing (in its specific region of validity X_i, $i = 1{:}s$) the present (k) sample of one selected variable y in terms the past of all the measured variables $\mathbf{x}(k) = [y(k-1)\ldots y(k-n_a)\quad \mathbf{u}'(k-1)\ldots\mathbf{u}'(k-n_b)]'$, via the corresponding set of parameters $\theta_i = [a_{i,1}\ldots a_{i,na}\ \mathbf{b}_{i,1}\ldots\mathbf{b}_{i,nb}\ \mathbf{f}_i]'$ whose joint identification together with the polyhedral partition $X_1{:}X_s$ of the polytope X is the goal of the approach proposed by Ferrari-Trecate *et al.*[8]

Piece-wise affine identification exploits *K-means*[17] clustering, thus associating data points in multivariable space in such a way to jointly determine a sequence of linear sub-models in the above form and their respective regions X_i of operation without even imposing continuity at each change in the derivative, and thus allowing to identify hybrid dynamic-logical processes. In order to obtain such a result, the five following steps are executed.

- *Step 1*: The proposed model is locally linear, thus small sets of data points close to each other likely belong to the same sub-model. Then, for each data point, a local set is built, collecting the selected point together with a given number of its neighbors (whose cardinality is one of the parameters of the algorithm). Each local set will be pure (like C_1 in the simple really piece-wise linear simulated example illustrated in Fig. 1) if made of points really belonging to the same single linear subsystem; otherwise, it is mixed, like C_2.
- *Step 2*: For each local dataset, a linear model is identified through usual least squares procedure. Pure sets belonging to the same sub-model give similar parameter sets, while mixed sets yield isolated vectors of coefficients, appearing as outliers in the space of the identified parameters shown in Fig. 2. If the signal to noise ratio is strong enough, and if there are not too many mixed sets (i.e., the number of data points is enough more than the number of sub-models to be identified, and the sampling is fair in every region), then the vectors will cluster in the parameter space around the values pertaining to each sub-model, apart from a few

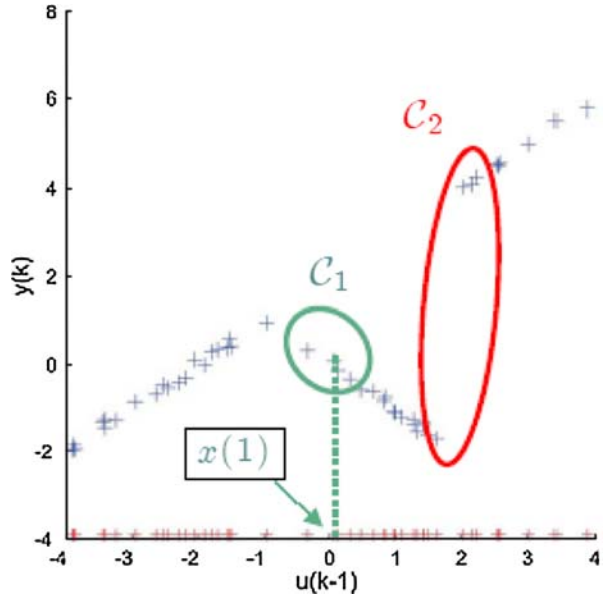

FIGURE 1. Examples of pure (C1) and mixed (C2) data sets from a simple example whose data are extracted form a simulated noisy process switching among different linear behaviors.

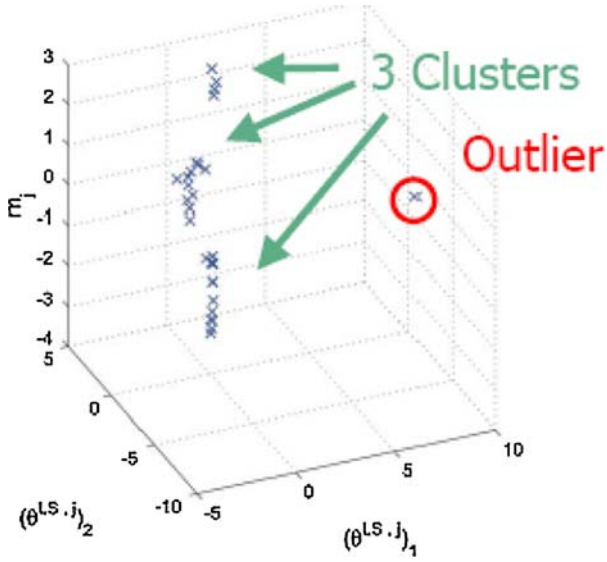

FIGURE 2. Clustering, in the space of the parameters defined in Eq. (1), of the sets identified for each of the data in Fig. 1.

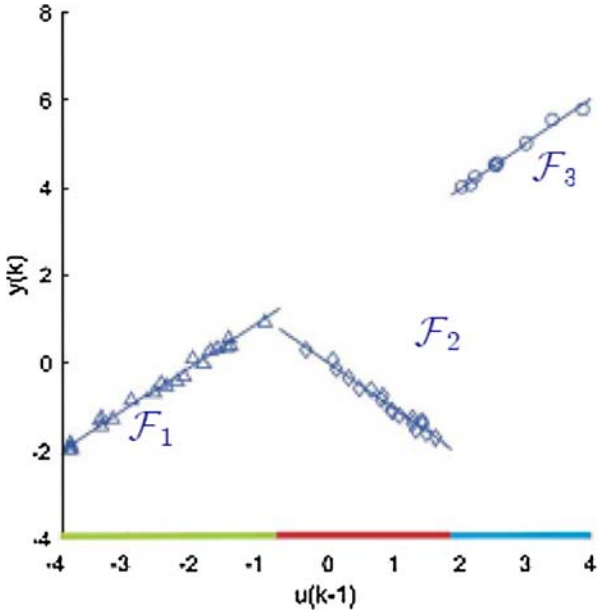

FIGURE 3. Classification of each data point of the example in Fig. 1: the estimated sub-model for each identified region is plotted.

outliers, as shown in Fig. 2 for the example in Fig. 1.

- *Step 3*: A modified version of the classical *K-means*, whose convergence is guaranteed in a finite number of steps,[8] takes into account the confidence on both pure and mixed local sets in order to cluster the parameter vectors.
- *Step 4*: Data points are then classified, each being a local dataset one-to-one related to its generating data point, which is thus classified according to the cluster to which its parameter vector belongs, as shown in Fig. 3 for the same example discussed in Figs. 1 and 2, where each sample has now a symbol according to his belonging.
- *Step 5*: Both the linear sub-models and their regions are estimated from the data in each subset, as also shown in Fig. 3. The coefficients are estimated via weighted least squares, taking into account the confidence measures. The shape of the polyhedral region characterizing the domain of each model may be obtained via linear support vector machines,[18] easily solved via linear/quadratic programming.

PWA IDENTIFICATION OF PATHO-PHYSIOLOGICAL SWITCHES

Some physiological processes are (naturally or intentionally) switching from an active to quiescent state. Such switches are considered and searched for via PWA approach in the following examples.

An example of physiological switching are hormone pulses in neuro-endocrinology, whose identification[16] is for instance important in growth and fertility diseases, as well as in doping assessment. As a first approximation of the secretion process, the involved gland switches from a quiescent status to a secretion status, which is in turn often approximated as a pulse. The smooth course of the secreted hormone measurable in the blood is thus the convolution of such asynchronous secreted pulses with the known hormone clearance. If one thus needs to reconstruct the firing pulsating time pattern of the secreting gland, often not directly measurable because of both accessibility problems (pituitary) and small amount of secretion at each pulse (leading to a great amount of false detection because of measurement errors), PWA provides an alternative easy approach, depicted in Fig. 4, with respect to the de-convolution recalled by DeNicolao *et al.*[6] and described in the literature cited there, in order to face the ill posed inverse problem to identify the originating neuro-endocrine secretion. In the upper panel of Fig. 4, the time course of a representative hormone blood concentration is depicted. Because of both blood (multi)-exponential clearance in time and concentration measurement errors, the originally pulsatile pituitary secretion is hardly detectable in the sampled series, while it would be useful to know pulse secretion events in order to assess the pituitary behavior with respect to the considered hormonal axis. Feeding the samples of such time course to the PWA identification tool, one identifies the belonging of each sampled point and its surroundings to either one among two possible sub-models, depicted in the lower panel at different ordinate for each point. The sub-model to which samples in the lower status in lower panel in Fig. 4 belong corresponds to the quiescent status of the gland, while the points in the upper status belong to pulses. Thus, lower panel reports the classification in the two behaviors obtained by the PWA application to the upper panel data, whose time course does estimate the subsequent switching among secretion and silence, thus filtering the noisy acquired data in order to recover the estimated secretion, whose computation is a known ill-posed problem, because of the inverse computation that a more traditional modeling approach would imply.

A different application, where switches are among a physiological and a possibly pathological state, is reported in Fig. 5, depicting the course of voluntary apneas alternated with regular breath in order to simulate possibly dangerous not voluntary apneas while sleeping, whose proper detection could trigger life support devices improving the quality of the sleep.

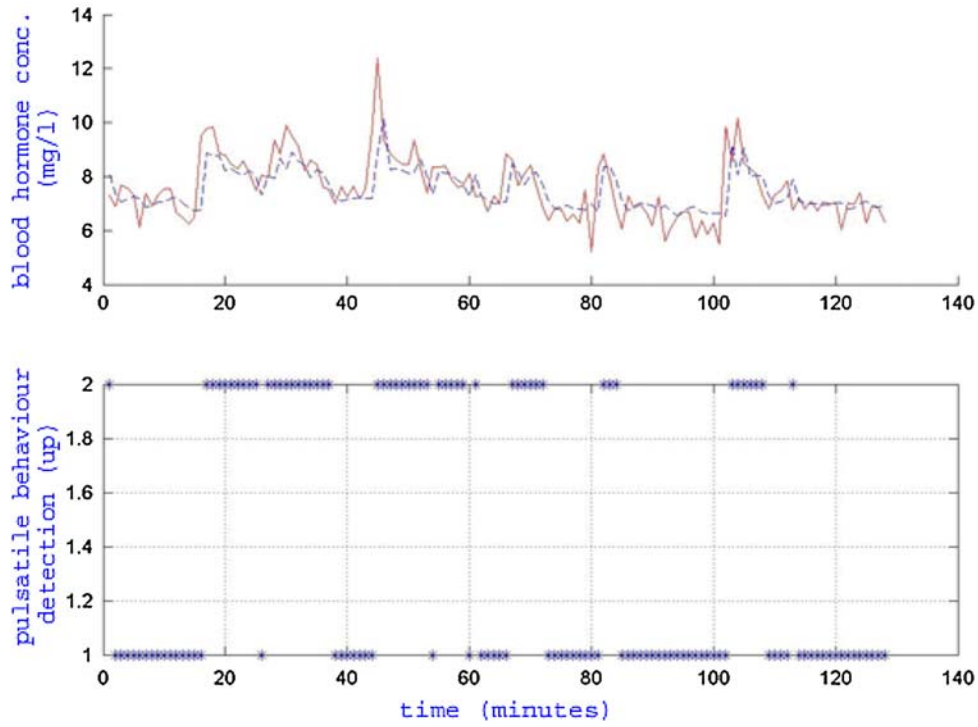

FIGURE 4. Time course of blood hormone concentration (upper panel) and PWA identification of originating pulses (upper state 2 in lower panel, where 1 is the quiescent state).

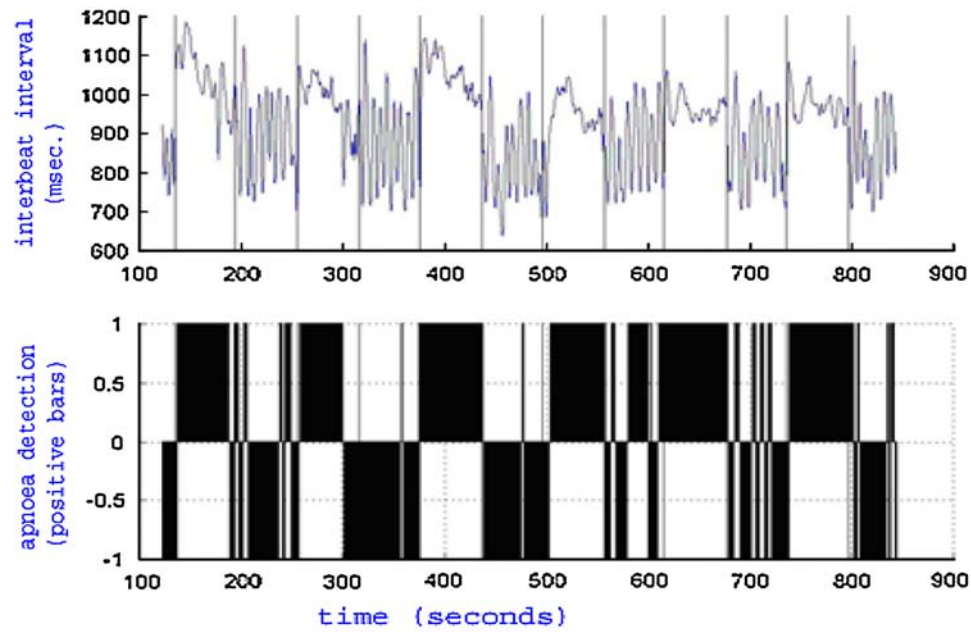

FIGURE 5. PWA apnea detector: the interval between each pair of subsequent heart beats is depicted in the upper plot over time, while the lower panel plots the results of the sliding window PWA identification (positive bars identify apnea detected intervals).

The identified variable in this approach is indeed the variation in inter-beat duration, whose dynamics is depicted in the upper panel, while in the lower panel the two identified states via PWA approach are plotted. The quite correct PWA capturing of apnea events induced by breath holding, roughly corresponding to

increasing inter-beat duration, is shown for a representative subject. Such a PWA tool may be used as apnea detector, feeding back to a controller that automatically supply more oxygen in a control closed loop, or at least awakes the subject in danger. By analyzing the same inter-beat variability with the help of PWA identification in other frameworks, a psychological stress indicator may be for instance inferred, complementing,[15] and used for bio-feedback.

DISCUSSION

Through the proposed approach, thus, the probably most fascinating human organ, namely the brain, may be at least indirectly addressed, whose study may also be undertaken today in the sophisticated frame of functional nuclear magnetic imaging[2] besides the simpler but still informative traditional way of EEG recording.[11] Multidimensional (three space variables plus time) images or multivariate time series provide a lot of raw data to be mined in order to understand which kind of activation is produced in correspondence with an event or a decision, which is indeed an intentional switch prone to be captured by PWA identification. Such multivariate analysis, possibly taking into account also the input stimulation, is useful in approaching not easy neurological tasks like modeling electro encephalographic coherence in Alzheimer's disease patients,[12] as well as nonlinear effects in muscle contraction,[14] could be complemented by piece-wise linear identification, even in time-varying contexts like epilepsy, thus leading to a possible implantable control device.

A brain-computer interfacing device may then be built that is able to reproduce one's intention to perform a move, not directly possible for the subject in some impaired physical conditions, and thus directly command a proper actuator. Cascading piece-wise affine identification to one of the recalled techniques able to prune redundant variables could improve the only partially successful approach based on artificial neural networks.[1] A simple drowsiness detector based on the EEG may be designed, similar to the apnea detector previously addressed, as well as a flexible anesthesia/hypotension level, or sleep apnea, detector, without needing a time-varying more precise, but more costly, identification.

A growing field of international relevance, in which the hybrid systems approaches are eager to provide a contribution, is the so-called field of systems biology—a feedback model of how proteins interact with each other and with nucleic acids within the cell—which is needed to better understand the control mechanisms of the cellular cycle, especially with respect to duplication (and, thus, cancer, when such mechanisms become out of control). Such understanding hopefully would drive to personal therapy, when everybody's gene expression will be correlated to the amount of corresponding proteins involved in the cellular cycle. Moreover, a new computational paradigm could arise by exploiting biological components like cells instead of the usual silicon hardware, thus overcoming some technological issues and possibly facilitating neuro-informatics. The study of systems biology, as a frontier edge of the larger field of bioinformatics, do also start from analyzing data from so-called micro-arrays. They are little standard chips, where thousands of gene expressions may be obtained from the same cell material in a while, thus providing a huge amount of data whose handling with the usual deterministic approaches is not conceivable, fault to obtain not significant synthetic information. Thus, matrices of as many subjects as available, possibly grouped in homogeneous categories for supervised training, each one carrying its luggage of thousands of gene expressions, are the natural input to approaches based on either Logical Networks[13] or Adaptive Bayesian Networks.[5] A desired output is to extract rules classifying; for instance, patients affected by different tumors, like in Garatti *et al.*,[9] and identifying their prognostic factors,[7] on the basis of a few identified genes, whose set is the candidate basis for the piece-wise linear model describing their complex interaction in such a particular class of subjects.[10]

The recalled approaches do constitute an easy way to face even complex identification problems from data, whose solution may contribute to understand and design control strategy, as also exemplified within the paper.

Industrial applications, of course, are not excluded from the field of possible interests. In Ferrari-Trecate *et al.*,[8] for instance, the classification and identification of the hysteretic dynamics of an industrial transformer for control purpose are performed via the piece-wise approach, with a much simpler cost, not really paid in significant reduction of performances with respect to the non-linear accurate physical modeling described in Bittanti *et al.*[4]

Among other possible approaches, remote sensing is keen to be helped by such an approach, useful for instance in detecting the shapes of the ceilings of the houses in urban environments and thus monitoring and control natural and manufactured changes.

ACKNOWLEDGMENT

Giancarlo Ferrari-Trecate is warmly thanked for being involved also in the physiological data analysis.

REFERENCES

[1]Babiloni, F., F. Carducci, S. Cerutti, D. Liberati, P. Rossini, A. Urbano, and C. Babiloni. Comparison between human and ANN detection of Laplacian-derived electroencephalographic activity related to unilateral voluntary movements. *Comput. Biomed. Res.* 33:59–74, 2000. doi:10.1006/cbmr.1999.1529.

[2]Baraldi, P., A. A. Manginelli, M. Maieron, D. Liberati, and C. A. Porro. An ARX model-based approach to trial by trial identification of fMRI-BOLD responses. *NeuroImage* 37:189–201, 2007. doi:10.1016/j.neuroimage.2007.02.045.

[3]Barron, A., J. Rissanen, and B. Yu. The minimum description length principle in coding and modelling. *IEEE Trans. Inf. Theory* 44:2743–2760, 1998. doi:10.1109/18.720554.

[4]Bittanti, S., F. A. Cuzzola, F. Lorito, and G. Poncia. Compensation of nonlinearities in a current transformer for the reconstruction of the primary current. *IEEE Trans. Control Syst. Technol* 9(4):565–573, 2001. doi:10.1109/87.930967.

[5]Bosin, A., N. Dessì, D. Liberati, and B. Pes. Learning Bayesian classifiers from gene-expression microarray data. Lecture Notes in Computer Science. New York: Springer-Verlag, vol. 3849, pp. 297–304. doi:10.1007/11676935_37.

[6]DeNicolao, G., D. Liberati, and A. Sartorio. Stimulated secretion of pituitary hormones in normal humans: a novel direct assessment from blood concentrations. *Ann. Biomed. Eng.* 28:1131–1145, 2000.

[7]Drago, G. P., E. Setti, L. Licitra, and D. Liberati. Forecasting the performance status of head and neck cancer patient treatment by an interval arithmetic pruned perceptron. *IEEE Trans. Biomed. Eng.* 49(8):782–787, 2002.

[8]Ferrari-Trecate, G., M. Muselli, D. Liberati, and M. Morari. A clustering technique for the identification of piecewise affine systems. *Automatica* 39:205–217, 2003. doi:10.1016/S0005-1098(02)00224-8.

[9]Garatti, S., S. Bittanti, D. Liberati, and P. Maffezzoli. An unsupervised clustering approach for leukemia classification based on DNA micro-arrays data. *Intell. Data Anal.* 11(2):175–188, 2007.

[10]de Jong, H., J.-L. Gouzé, C. Hernandez, M. Page, T. Sari, and J. Geiselmann. Qualitative simulation of genetic regulatory networks using piecewise-linear models. *Bull. Math. Biol.* 66(2):301–340, 2004. doi:10.1016/j.bulm.2003.08.010.

[11]Liberati, D., M. Cursi, T. Locatelli, G. Comi, and S. Cerutti. Total and partial coherence of spontaneous and evoked EEG by means of multi-variable autoregressive processing. *Med. Biol. Eng. Comput.* 35(2):124–130, 1997. doi:10.1007/BF02534142.

[12]Locatelli, T., M. Cursi, D. Liberati, M. Franceschi, and G. Comi. EEG coherence in Alzheimer's disease. *Electroencephalogr. Clin. Neurophysiol.* 106(3):229–237, 1998. doi:10.1016/S0013-4694(97)00129-6.

[13]Muselli, M., and D. Liberati. Binary rule generation via hamming clustering. *IEEE Trans. Knowl. Data Eng.* 14(6):1258–1268, 2002. doi:10.1109/TKDE.2002.1047766.

[14]Orizio, C., D. Liberati, C. Locatelli, D. DeGrandis, and A. Veicsteinas. Surface mechanomyogram reflects muscle fibres twitches summation. *J. Biomech.* 29(4):475–481, 1996. doi:10.1016/0021-9290(95)00063-1.

[15]Pagani, M., G. Mazzuero, A. Ferrari, D. Liberati, S. Cerutti, D. Vaitl, L. Tavazzi, and A. Malliani. Sympathovagal interaction during mental stress: a study employing spectral analysis of heart rate variability in healthy controls and in patients with a prior myocardial infarction. *Circulation 83*(4), II43–II51, 1991.

[16]Sartorio, A., G. De Nicolao, and D. Liberati. An improved computational method to assess pituitary responsiveness to secretagogue stimuli. *Eur. J. Endocrinol.* 147(3):323–332, 2002. doi:10.1530/eje.0.1470323.

[17]Selim, S. Z., and M. A. Ismail. *K-means*-type algorithms: a generalized convergence theorem and characterization of local optimality. *IEEE Trans. Pattern Anal. Mach. Intell.* 6(1):81–86, 1984. doi:10.1109/TPAMI.1984.4767478.

[18]Vapnik, V. Statistical Learning Theory. New York: Wiley, 1998.