Available online at www.sciencedirect.com

ScienceDirect

Journal of the Franklin Institute 344 (2007) 813–829

Journal *of* The Franklin Institute

www.elsevier.com/locate/jfranklin

Biorthogonal wavelet trees in the classification of embedded signal classes for intelligent sensors using machine learning applications

Paolo Mercorelli*

Department of Vehicles, Production and Process Engineering, University of Applied Sciences Wolfsburg, Robert-Koch-Platz 8-a., 38440 Wolfsburg, Germany

Received 10 October 2006; accepted 11 October 2006

Abstract

The paper deals with a method of constructing orthonormal bases of coordinates which maximize, through redundant dictionaries (frames) of biorthogonal bases, a class separability index or *distances* among classes. The method proposes an algorithm which consists of biorthogonal expansions over two *redundant dictionaries*. *Embedded classes* are often present in multiclassification problems. It is shown how the biorthogonality of the expansion can really help to construct a coordinate system which characterizes the classes. The algorithm is created for training wavelet networks in order to provide an efficient coordinate system maximizing the *Cross Entropy function* between two complementary classes. Sine and cosine wavelet packets are basis functions of the network. Thanks to their packet structure, once selected the depth of the tree, an adaptive number of basis functions is automatically chosen. The algorithm is also able to carry out centering and dilation of the basis functions in an adaptive way. The algorithm works with a preliminary extracted feature through shrinkage technique in order to reduce the dimensionality of the problem. In particular, our attention is pointed out for time-frequency monitoring, detection and classification of transients in rail vehicle systems and the outlier problem. In the former case the goal is to distinguish transients as *inrush current* and *no inrush current* and a further distinction between the two complementary classes: *dangerous inrush current* and *no dangerous inrush current*. The proposed algorithm is used on line in order to recognize the dangerous transients in real time and thus shut-down the vehicle. The algorithm can also be used in a general application of the outlier detection. A similar structure is used in developed algorithms which are currently integrated in the inferential modeling platform of the unit responsible for Advanced Control and Simulation Solutions within ABB's (Asea Brown Boveri) industry division. It is shown how impressive and rapid performances are achieved with a limited

*Tel.: +49 5361 831615; fax: +49 5361 831602.

E-mail addresses: p.mercorelli@fh-wolfsburg.de, mercorelli@ing.unisi.it (P. Mercorelli).

doi:10.1016/j.jfranklin.2006.10.003

number of wavelets and few iterations. Real applications using real measured data are included to illustrate and analyze the effectiveness of the proposed method.

Keywords: Machine learning; Signal classification; Trigonometric bases; Wavelet packets; Wavelet networks

1. Introduction and problem definition

Extracting relevant features from signals or images is an important process for data analysis. In particular, when the problem consists of classifying signals into known categories, the approach is similar to a nonparametric estimation. Detailed works are for instance [1,2]. This paper proposes a method, in learning machines, for the detection and classification of time-frequency phenomena by using wavelet packets in wavelet networks with biorthogonal functions. Specifically the task is to classify transient harmonics as they occur in electrical power systems [3]. Literature has indicated wavelets and wavelet networks as a promising approach for off line analysis and classification. Furthermore, a recent paper in this direction [4] developed an efficient algorithm which detects and classifies transient harmonic phenomena as *inrush current* and which presents a unimodal distribution. This paper tries to extend the previous results as in [4], by making several structural modifications to the training algorithm, to the class of *embedded phenomena* with which, very often, it can be modeled into a classification system. In other words it is possible to think of the embedded phenomena as belonging to the following structured classes.

Definition 1. Given a set $\mathscr{Y}$ of $(L+1)$ classes of signals: $\mathscr{Y} = \{\mathscr{C}_1, \mathscr{C}_2, \ldots, \mathscr{C}_L, \mathscr{C}_C\}$, where L are known classes and $\mathscr{C}_C$ is the class defined as complement to the class $\mathscr{C}_L$. Let set $\mathscr{Y}$ be a set of *embedded classes* if the following relationships hold:

$$\mathscr{C}_L \cap \mathscr{C}_C = \{0\} \tag{1}$$

and

$$\begin{aligned} &\mathscr{C}_1 \cap \mathscr{C}_2 = \{0\}; \quad \mathscr{C}_1 \cup \mathscr{C}_2 = \mathscr{C}_3;\\ &\mathscr{C}_3 \cap \mathscr{C}_4 = \{0\}; \quad \mathscr{C}_3 \cup \mathscr{C}_4 = \mathscr{C}_5;\\ &\qquad\qquad \vdots \\ &\mathscr{C}_{(L-2)} \cap \mathscr{C}_{(L-1)} = \{0\}; \quad \mathscr{C}_{(L-2)} \cup \mathscr{C}_{(L-1)} = \mathscr{C}_L. \end{aligned} \tag{2}$$

In our former task the embedded classes could be seen in the following way. Let $L = 3$, and let $\mathscr{C}_1 = \{$*No dangerous inrush current*$\}$ its relative complementary class such as $\mathscr{C}_2 = \{$*Dangerous inrush current*$\}$ and thus at the end $\mathscr{C}_3 = \{$*Inrush current*$\}$ with its relative complementary class $\mathscr{C}_C = \{$*No inrush current*$\}$. It holds: $\mathscr{C}_3 \cap \mathscr{C}_C = \{0\}$ and $\mathscr{C}_1 \cup \mathscr{C}_2 = \mathscr{C}_3$. The problem is to classify the inrush current which has, for instance, an energy which exceeds several limits in several particular bands of frequency. In [4], the proposed algorithm looked for the best basis (Shannon Entropy function) over a dual frame (smooth trigonometric sine/cosine wavelet functions) and through that a dual coordinate system can be effectively determined using a particular training algorithm for wavelet networks. In [4] the task was *to detect and classify inrush current*. In this paper the problem looks to be a