Chapter 10
Nonstationary Signal Analysis with Kernel Machines

Paul Honeine
Institut Charles Delaunay, France

Cédric Richard
Institut Charles Delaunay, France

Patrick Flandrin
Ecole Normale Supérieure de Lyon, France

ABSTRACT

This chapter introduces machine learning for nonstationary signal analysis and classification. It argues that machine learning based on the theory of reproducing kernels can be extended to nonstationary signal analysis and classification. The authors show that some specific reproducing kernels allow pattern recognition algorithm to operate in the time-frequency domain. Furthermore, the authors study the selection of the reproducing kernel for a nonstationary signal classification problem. For this purpose, the kernel-target alignment as a selection criterion is investigated, yielding the optimal time-frequency representation for a given classification problem. These links offer new perspectives in the field of nonstationary signal analysis, which can benefit from recent developments of statistical learning theory and pattern recognition.

INTRODUCTION

Time-frequency and time-scale distributions have become increasingly popular tools for analysis and processing of nonstationary signals. These tools map a one-dimensional signal into a two-dimensional distribution, a function of both time and frequency. Such joint description reveals the time-varying frequency content of nonstationary signals, unlike classical spectral analysis techniques *a la Fourier*. Over the years, a large variety of classes of time-frequency distributions have been proposed to explain the diversity of the treated problems. Linear and quadratic distributions have been extensively studied, and among them Cohen's class of time-frequency distributions as (quasi) energy distribution jointly

DOI: 10.4018/978-1-60566-766-9.ch010

in time and frequency, and most notably the Wigner-Ville distribution, see for instance (Cohen, 1989; Flandrin, 1999; Auger & Hlawatsch, 2008). From these, one can choose the optimal representation for the problem under investigation, such as increasing the representation immunity to noise and interference components (Auger & Flandrin, 1995, Baraniuk & Jones, 1993), or selecting the best class of representations for a given decision problem (Heitz, 1995; Till & Rudolph, 2000; Davy, Doncarly, & Boudreaux-Bartels, 2001).

Over the last decade, multiple analysis and classification algorithms based on the theory of reproducing kernel Hilbert space (RKHS) have gained wide popularity. These techniques take advantage of the so-called *kernel trick*, which allows construction of nonlinear techniques based on linear ones. Initiated by state-of-the-art support vector machines (SVM) for classification and regression (Vapnik, 1995), the most popular ones include the nonlinear generalization of principal component analysis or kernel-PCA (Schölkopf, Smola, & Müller, 1998), and nonlinear Fisher's discriminant analysis or kernel-FDA (Mika, Rätsch, Weston, Schölkopf, & Müller, 1999); see (Shawe-Taylor & Cristianini, 2004) for a survey of kernel machines. Kernel machines are computationally attractive, with outstanding performance, validated theoretically by the statistical learning theory (Vapnik, 1995; Cucker & Smale, 2002). Despite these advances, nonstationary signal analysis and classification still has not benefited from these developments, although such techniques have been brought to the attention of the signal processing community. Few work combine kernel machines and time-frequency analysis, of these Davy *et al.* (2002) apply the SVM algorithm for classification with a reproducing kernel expressed in the time-frequency domain. More recently, Honeine *et al.* (2007) applied a large panel of kernel machines for nonstationary signal analysis and classification, while in Honeine *et al.* (2006) and in Honeine and Richard (2007) the optimality of the representation space is treated.

This chapter shows how the most effective and innovative kernel machines can be configured, with a proper choice of reproducing kernel, to operate in the time-frequency domain. Further, this approach is extended to the selection of the optimal time-frequency domain for a given classification task. For this purpose, the strength of the kernel-target alignment criterion is investigated for time-frequency distributions. The performance of the proposed approach is illustrated with simulation results. But before, a brief review of the principal elements of the theory behind kernel machines is presented.

RKHS AND KERNEL MACHINES: A BRIEF REVIEW

The theory behind RKHS serves as a foundation of the kernel machines. The main building blocks of these statistical learning algorithms are the *kernel trick* and the Representer Theorem. In this section, these concepts are presented succinctly, after a short introduction on reproducing kernels.

Reproducing Kernels and RKHS

Let X be a subspace of $L_2(C)$ the space of finite-energy complex signals, equipped with the usual inner product defined by $\langle x_i, x_j \rangle = \int_t x_i(t)\, x_j^*(t)\, dt$ and its corresponding norm, where $x_j^*(t)$ denotes the complex conjugate of the signal $x_j(t)$. A kernel is a function $\kappa(x_i, x_j)$ from X×X to C, with Hermitian symmetry. The basic concept of reproducing kernels is described by the following two definitions (Aronszajn, 1950).

Related Content

Sustainable Supplier's Management Using Differential Evolution
Sunil Kumar Jauharand Millie Pant (2016). *Problem Solving and Uncertainty Modeling through Optimization and Soft Computing Applications (pp. 239-263).*
www.irma-international.org/chapter/sustainable-suppliers-management-using-differential-evolution/147094

Power Pell Sequences, Some Periodic Relations of These Sequences, and a Cryptographic Application With Power Pell Sequences
Çala Çelemolu, Selime Beyza Özçevikand enol Eren (2022). *Applications of Computational Science in Artificial Intelligence (pp. 21-43).*
www.irma-international.org/chapter/power-pell-sequences-some-periodic-relations-of-these-sequences-and-a-cryptographic-application-with-power-pell-sequences/302060

EEG Analysis to Decode Tactile Sensory Perception Using Neural Techniques
Anuradha Sahaand Amit Konar (2018). *Applied Computational Intelligence and Soft Computing in Engineering (pp. 178-203).*
www.irma-international.org/chapter/eeg-analysis-to-decode-tactile-sensory-perception-using-neural-techniques/189321

Learning and Prediction of Complex Molecular Structure-Property Relationships: Issues and Strategies for Modeling Intestinal Absorption for Drug Discovery
Rahul Singh (2012). *Machine Learning: Concepts, Methodologies, Tools and Applications (pp. 1482-1498).*
www.irma-international.org/chapter/learning-prediction-complex-molecular-structure/56208

Optimization of the Wake-Up Scheduling Using a Hybrid of Memetic and Tabu Search Algorithms for 3D-Wireless Sensor Networks
Vrajesh Kumar Chawraand Govind P. Gupta (2022). *International Journal of Software Science and Computational Intelligence (pp. 1-18).*
www.irma-international.org/article/optimization-of-the-wake-up-scheduling-using-a-hybrid-of-memetic-and-tabu-search-algorithms-for-3d-wireless-sensor-networks/300359