Information Sciences 283 (2014) 153–164

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Best basis for joint representation: The median of marginal best bases for low cost information exchanges in distributed signal representation

Abdourrahmane M. Atto *, Kavé Salamatian, Philippe Bolon

University of Savoie, Polytech Annecy-chambéry, LISTIC, Laboratory of Computer Science, Systems, Information and Knowledge Processing, BP 80439, F-74944 Annecy-le-Vieux Cedex, France

ARTICLE INFO

Article history:
Received 10 October 2013
Received in revised form 17 June 2014
Accepted 26 June 2014
Available online 3 July 2014

Keywords:
Distributed signal
Best joint basis
Information fusion
Wavelet

ABSTRACT

The paper addresses the selection of the best representations for distributed and/or dependent signals. Given an indexed tree structured library of bases and a semi-collaborative distribution scheme associated with minimum information exchange (emission and reception of one single index corresponding to a marginal best basis), the paper proposes the median basis computed on a set of best marginal bases for joint representation or fusion of distributed/dependent signals. The paper provides algorithms for computing this median basis with respect to standard tree structured libraries of bases such as wavelet packet bases or cosine trees. These algorithms are effective when an additive information cost is under consideration. Experimental results performed on distributed signal compression confirms worthwhile properties for the median of marginal best bases with respect to the ideal best joint basis, the latter being underdetermined in practice, except when a full collaboration scheme is under consideration.

1. Introduction

Among the functional representations one can associate with a *regular* or piecewise regular deterministic signal, low cost information representations are important for many applications involving compression, coding, estimation, dimensionality reduction, *etc.* In general finding a relevant representation for a whole class of signals is intricate, especially when signals under consideration are impacted by uncertainties/imprecisions inherent to the measurement process and/or specific noise emanating from the acquisition system or external disturbances.

Adaptive and/or fuzzy approaches have shown to be relevant for joint analysis and processing of such class of signals. For instance, in [12] a statistical model associating fuzzy regression, nearest neighbor matching, and neural networks has been proposed for predicting the demand of natural gas by using heterogeneous rooftop unit wireless sensors; in [13] a fuzzy multi-sensor data fusion and a fuzzy Kalman feedback are used for fault detection and effective risk reduction for an integrated vehicle health maintenance system; the analysis of a neuro-fuzzy system involving adaptive wavelet activation that depends on the input signal characteristics is described in [3]; the authors of [10] show that genetic algorithms based on lifting (and thus adaptive) wavelet transforms enables relevant source separation for wide band signals while diminishing

* Corresponding author.
E-mail addresses: Abdourrahmane.Atto@univ-savoie.fr (A.M. Atto), Kave.Salamatian@univ-savoie.fr (K. Salamatian), Philippe.Bolon@univ-savoie.fr (P. Bolon).

http://dx.doi.org/10.1016/j.ins.2014.06.040
0020-0255/

different types of noises; in [16] the correlation structure is used to improve the estimation accuracy of highly correlated measurements performed in multi-sensor systems.

In this paper, we analyze a distributed set of signals by using a library of wavelet functions. In contrast to [3,10], this paper does not involve adaptive prediction and updating of wavelet coefficients (wavelet lifting). We first derive a finite set of relevant wavelet base for representing a distributed set of signals through a 'lower' and 'upper' wavelet basis, delimiting the set of bases-of-interest (fuzzy functional set). A functional ordering of the bases-of-interest is then proposed for selecting the best joint-basis for distributed compression or coding.

Refs. [17,5–7,11] provide concentration norms and sparsity information costs for best basis selection with respect to one signal observation (*marginal best basis* when considering a distributed set of dependent signals). However, for a distributed acquisition system involving dependent and non-stationary signals, finding a best basis from a joint criterion is not an easy extension of the single signal acquisition case. Given a joint information criterion, the best joint basis cannot be computed without a full collaboration between sensors or gathering of all the data at a central node, whereas both situations are undesirable due to their consequence on sensor architectures and their energy consumptions [9]. In this respect and in order to approach a locally optimal solution, some references such as [1] have investigated semi-collaborative distributions schemes consisting of recursive implementations of the Karhunen–Loève transform. On the other hand, [3,10,14,4,15] have considered wavelet node splitting subject to prediction and updating stages and conditionally with respect to pre-specified collaboration schemes.

It is worth recalling that the above alternatives are with high computational complexities when the number of sensors/signals is large. In addition, these strategies do not guarantee a convergence to the ideal best joint basis even when the number of recursive operations is significant.

The motivation of the present work is to seek, from the sole knowledge of the marginal best bases (best bases at each individual sensor levels), a basis approximating the joint best basis, which is unknown and undetermined in case of non-collaborative distribution schemes. We show that, on tree structured libraries of functional bases, the median of marginal best bases is a basis with relevant properties for joint representation. In addition, the tree structuring makes the computation of the infimum and supremum bases possible, the latter being useful for evaluating the dispersion of a set of marginal best bases. The paper provides theoretical concepts and algorithmic tools that make the computation of the median, infimum and supremum of a best basis set associated with an additive information cost possible over a tree structured library.

The paper begins by providing the context of best basis selection for distributed signals (Section 2). Then it focuses on defining algebra on tree structured libraries (Section 3). From this algebra, the paper derives algorithms for computing the median, infimum and supremum of a set of bases-of-interests (Section 3). The paper then highlights the relevance of the sample median basis for joint representation when the observation amounts solely to marginal basis consideration (Section 4). The paper finally concludes by discussing some issues concerning the use and interpretation of median, infimum and supremum bases (Section 5).

2. Preliminary notation and issues

2.1. Context

Let us consider distributed compression or distributed fusion of a set of signals delivered by K sensors, *i.e.*, K observations $\mathbf{y_k}$ (partial "views") of a "big" signal s. These observations are available through a model of the form:

$$\mathbf{y_k} = \Theta_k(\mathbf{s}, \xi_k), \tag{1}$$

where $\Theta_k,\ \xi_k$ for $k = 1, 2, \ldots, K$, are respectively operators and noise relating the specificities of the sources/sensors.

Operator Θ_k can be additive (signal $\mathbf{s}$ observed in presence of additive noise ξ_1), multiplicative (acquisition systems using coherent radiations), convolutive (transfer function involved in some imaging systems) or masking (missing data inducing a partial loss of information or a partial observation of a whole phenomenon), *etc.*

Fig. 1 provides an illustration of the model of Eq. (1) where operator Θ_k is additive with respect to variables $\mathbf{s}$ and ξ_k, and masking (has a limited access to the whole signal s):

$$\mathbf{y_k} = \mathbf{s}\mathbb{1}_{\Delta_k} + \xi_k, \tag{2}$$

where the intervals $(\Delta_k)_{k=1,2,\ldots,K}$ involved in Eq. (2) overlap, yielding dependent observations $(\mathbf{y_k})_{k=1,2,\ldots,K}$. Since multiplicative or convolutive operators can be written in the form of Eq. (2) with appropriate transforms, we will use the distributed system given by the model of Eq. (2) in the following, for the sake of simplicity of presentation. We will moreover use for convenience, the notation $\mathbf{s}_k = \mathbf{s}\mathbb{1}_{\Delta_k}$.

Let us assume that there exists a relevant library of functional bases $\mathcal{B} = \{\mathbf{B}_\ell,\ \ell = 1, 2, \ldots, L\}$ for representing the signal. In order to avoid any confusion when several bases are under consideration, we will denote the representation $\mathbf{B}_\ell[\mathbf{s}_k]$ of signal $\mathbf{s}_k$ on a basis $\mathbf{B}_\ell \in \mathcal{B}$ by $\mathbf{s}_k^{\mathbf{B}_\ell}$.

We evaluate the relevance of representations in different bases $(\mathbf{s}_k^{\mathbf{B}_\ell})_{\ell=1,2,\ldots,L}$ in the particular context of the targeted application, *e.g.*, compression, coding or information fusion. For this purpose, we use an information cost function that is attached to every representation. In the following Section 3, there will be no need to detail the particular information cost function, we