ARTICLE IN PRESS YACHA:1082

Appl. Comput. Harmon. Anal. • • • (• • • •) • • •–• • •

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

www.elsevier.com/locate/acha

Sparse representation on graphs by tight wavelet frames and applications [☆]

Bin Dong [1]

Beijing International Center for Mathematical Research (BICMR), Peking University, Beijing, 100871, China

ARTICLE INFO

Article history:
Received 13 November 2014
Received in revised form 9 September 2015
Accepted 18 September 2015
Available online xxxx
Communicated by Stephane Jaffard

Keywords:
Tight wavelet frames
Sparse approximation on graphs
Spectral graph theory
Big data
Graph clustering

ABSTRACT

In this paper, we introduce a new (constructive) characterization of tight wavelet frames on non-flat domains in both continuum setting, i.e. on manifolds, and discrete setting, i.e. on graphs; we discuss how fast tight wavelet frame transforms can be computed and how they can be effectively used to process graph data. We start with defining the quasi-affine systems on a given manifold $\mathcal{M}$. The quasi-affine system is formed by generalized dilations and shifts of a finite collection of wavelet functions $\Psi := \{\psi_j : 1 \le j \le r\} \subset L_2(\mathbb{R})$. We further require that ψ_j is generated by some refinable function ϕ with mask a_j. We present the condition needed for the masks $\{a_j : 0 \le j \le r\}$, as well as regularity conditions needed for ϕ and ψ_j, so that the associated quasi-affine system generated by Ψ is a tight frame for $L_2(\mathcal{M})$. The condition needed for the masks is a simple set of algebraic equations which are not only easy to verify for a given set of masks $\{a_j\}$, but also make the construction of $\{a_j\}$ entirely painless. Then, we discuss how the transition from the continuum (manifolds) to the discrete setting (graphs) can be naturally done. In order for the proposed discrete tight wavelet frame transforms to be useful in applications, we show how the transforms can be computed efficiently and accurately by proposing the fast tight wavelet frame transforms for graph data (WFTG). Finally, we consider two specific applications of the proposed WFTG: graph data denoising and semi-supervised clustering. Utilizing the sparse representation provided by the WFTG, we propose ℓ_1-norm based optimization models on graphs for denoising and semi-supervised clustering. On one hand, our numerical results show significant advantage of the WFTG over the spectral graph wavelet transform (SGWT) by [1] for both applications. On the other hand, numerical experiments on two real data sets show that the proposed semi-supervised clustering model using the WFTG is overall competitive with the state-of-the-art methods developed in the literature of high-dimensional data classification, and is superior to some of these methods.

☆ This work was supported in part by NSF DMS-1418772.

E-mail address: dongbin@math.pku.edu.cn.

[1] The author is also supported by the Thousand Talents Plan of China.

http://dx.doi.org/10.1016/j.acha.2015.09.005
1063-5203/

1. Introduction

In recent years, we are experiencing rapid advances in information and computer technology, which contribute greatly to the exponential growth of data. To properly handle, process and analyze such huge and often unstructured data sets, sophisticated mathematical tools and efficient computing methods need to be developed. Such huge data sets, commonly referred to as "big data", are generally modeled as huge graphs living in very high dimensional spaces. Graphs are commonly understood as a certain discretization or a random sample from some smooth Riemannian manifold [2–6]. To understand and analyze graphs and data on graphs (will be called graph data), the graph Laplacian is widely used to reveal the geometric properties of the graph and plays an important role in many applications such as graph clustering.

In signal and image processing, many methods are transform based. Sparsity of the signal/image to be recovered under a certain transform is the key to the success of many existing algorithms. One of the successful examples is the wavelet frame transform, especially the tight wavelet frame transform [7–19]. The power of tight wavelet frames lies in their ability to sparsely approximate piecewise smooth functions and the existence of fast decomposition and reconstruction algorithms. Recently, geometric properties of tight wavelet frames were discovered by connecting them to differential operators under variational and PDE frameworks [17–19].

The success of wavelet frames for data defined on flat domains motivates much research on generalizing wavelets and wavelet frames to curved, irregular and unstructured domains. In this paper, we introduce a (constructive) characterization of tight wavelet frames on non-flat domains in both continuum (on manifolds) and discrete (on graphs) setting, discuss how fast tight wavelet frame transforms can be computed and how they can be effectively used to process and analyze graph data. The basic idea is to understand eigenfunctions of Laplace–Beltrami operator (graph Laplacian in discrete setting) as Fourier basis on manifolds (graphs in discrete setting) and the associated eigenvalues as frequency components. This idea was used earlier by [1] in the discrete setting. In this paper, we further observe that Quasi-affine systems generated by dilations and shifts of wavelet functions can be defined on manifolds. When the elements in the quasi-affine system are generated from a refinable function, the transition from continuum (manifolds) to discrete (graphs) setting can be done very naturally. More importantly, such consideration makes the construction of various types of tight wavelet frames on manifolds/graphs totally painless, and it ensures the existence of fast decomposition and reconstruction algorithms which is crucial to many applications.

Given a compact and connected Riemannian manifold $(\mathcal{M}, g)$, denote $L_2(\mathcal{M})$ the space of square integrable functions on $\mathcal{M}$. We start with defining the quasi-affine system on $\mathcal{M}$. The quasi-affine system is formed by generalized dilations and shifts of a finite collection of wavelet functions $\Psi := \{\psi_j : 1 \leq j \leq r\} \subset L_2(\mathbb{R})$. We further restrict our consideration of Ψ to those that are generated by a set of masks $\{a_j : 0 \leq j \leq r\} \subset \ell_2(\mathbb{Z})$. Then, we present the condition needed for the masks $\{a_j : 0 \leq j \leq r\}$ (i.e. equation (2.15)) so that the associated quasi-affine system generated by Ψ is a tight frame for $L_2(\mathcal{M})$ (Theorem 2.1). The condition on the masks is a simple set of algebraic equations which are not only easy to verify for a given set of masks $\{a_j\}$, but also make the construction of $\{a_j\}$ painless. In particular, we show that under suitable conditions, the quasi-affine system on $\mathcal{M}$ generated by any set of the framelets constructed from the unitary extension principle [20] on $\mathbb{R}$ is a tight frame for $L_2(\mathcal{M})$ (Corollary 2.1). In addition, many masks constructed in [18] satisfy the conditions (2.15) in Theorem 2.1 as well, although they may not satisfy the unitary extension principle. Therefore, Theorem 2.1 is not only a rather generic characterization of tight wavelet frames for $L_2(\mathcal{M})$, it also provides a simple way of verifying and constructing various types of tight wavelet frame systems on $L_2(\mathcal{M})$.

Thanks to the aforementioned special consideration on Ψ, i.e. associating Ψ with a set of masks $\{a_j\}$, we discuss how the transition from the continuum (manifolds) to the discrete setting (graphs) can be naturally done. We show that inner products with wavelet frame functions on manifolds can be approximated in the discrete setting by "filtering" with the associated masks on graphs. This leads to multi-level discrete tight

Please cite this article in press as: B. Dong, Sparse representation on graphs by tight wavelet frames and applications, Appl. Comput. Harmon. Anal. (2015), http://dx.doi.org/10.1016/j.acha.2015.09.005