GREEDY DICTIONARY LEARNING METHODS FOR SPARSE REPRESENTATION OF SIGNALS

A THESIS

submitted by

VINAYAK ABROL

for the award of the degree

of

DOCTOR OF PHILOSOPHY

SCHOOL OF COMPUTING AND ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY MANDI

September, 2017

ॐ

The love of a devotee freezes the principle "ब्रह्म" in the form of his choice, say it Shiva or Durga. Not a religion, this is the way to quality life and inner peace.

This is democracy, spiritual democracy !
This is HINDUISM

– Adopted from the Hindu contribution by Jay Lakhani
towards BBC documentary "The Story of God"

स्वच्छ भारत
एक कदम स्वच्छता की ओर

To My Parents

Madhu Abrol and Yog Raj Abrol

And To My Brother

Keshav Abrol

Declaration

I hereby declare that the entire work embodied in this thesis is the result of investigations carried out by me in the **School of Computing and Electrical Engineering, Indian Institute of Technology Mandi**, with the supervision of **Dr. Anil Kumar Sao**, and that it has not been submitted elsewhere for any degree or diploma. In keeping with the general practice, due acknowledgments have been made wherever the work described is based on finding of other investigators.

Mandi, 175005

Date:

Vinayak Abrol

THESIS CERTIFICATE

This is to certify that the thesis titled **GREEDY DICTIONARY LEARNING METHODS FOR SPARSE REPRESENTATION OF SIGNALS**, submitted by **Vinayak Abrol**, to the Indian Institute of Technology, Mandi, for the award of the degree of **Doctor of Philosophy**, is a bonafide record of the research work done by him with my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other institute or university for the award of any degree or diploma.

Mandi, 175005

Date:

Dr. Anil Kumar Sao

(Ph.D Supervisor)

Acknowledgments

The best part of my days at IIT Mandi has undoubtedly been getting to meet and work with so many amazing people. I would like to thank faculty members: Dr. A. Bhavsar, Dr. Dileep A. D. and Dr. R. M. Rameshan, and my collaborators: Dr. S. Faghihroohi, Dr. A. Kassim (NUS Singapore) and Dr. B. Biswal (NJIT, USA), for their guidance. Importantly, my profound gratitude to TCS India for funding my Ph.D work along with IIT Mandi, DST-SERB, and IEEE SPS for supporting my international conference travels.

I take this opportunity to express my deep regards to my supervisor Dr. A. K. Sao for his exemplary guidance, monitoring and constant encouragement throughout the course of this research work. Regardless of the growing responsibility towards his other students and due to those endless administrative meetings for being the chairperson, he was always willing to take time for adding new ideas to my work. I am also very grateful to my doctoral committee for their review and suggestions: to Dr. Samar Agnihotri for some very challenging but motivating discussions; to Dr. Padmanabhan Rajan for direct or indirect sharp technical insights or for warm words of encouragement; to Dr. Sayed Abbas for encouraging scientific rigor while promoting a relaxed and helpful environment.

At IIT Mandi, I could not have asked for a more fun or interesting group of people to work with: Pulkit, Srimanta, Pravindra, Nivedita, Vibha, Shaifu, Seema, Priyanka, Anshul, Kartik, Prabjhot, and many more. Thank you all for making MAS Lab a place I truly enjoyed coming to work. I will fondly remember our favorite pastimes: competing in Candy Crush, discussing religious and political views, long lunches, evening tea conversations, finding ways to skip our weekly lab meetings, etc. Finally, I would like to thank my parents who supported me in every phase of life and encouraged me to achieve my goals.

Vinayak Abrol

Abstract

In the framework of sparse representation, signals are expressed as a linear combinations of few atoms from a resource called "dictionary". Sparse representations has received a lot of importance in various signal processing applications, and the chances of getting the sparsest representation is better in an overcomplete dictionary, where the number of atoms are more than the number of signal dimension. In the last decade, various dictionary learning (DL) algorithms have been developed, with varying algorithmic complexity and success rate. However, the thumb rule "*sparser the representation, better the dictionary*", requires a dictionary which is in general a dense matrix. Thus, manipulation of the dictionary can be computationally costly both at the DL stage and later in its use for signal processing applications. This thesis work is an attempt to address this problem and is mainly focused towards an alternative and efficient DL approach for sparse representation of signals and its application to several signal processing applications. In particular, this thesis explores the framework of learning a factorisable dictionary where a dictionary can be expressed as a product of matrices which are sparse or have an analytical form, and the individual factors can be chosen specific for a particular application. This makes the dictionary intrinsically fast to learn or manipulate and it requires less storage.

In order to speed up the DL process, the proposed approach performs dictionary preconditioning by modifying the underlying objective function in DL problem to a transformed space, which results in operations with only sparse matrices. This approach is further extended to the cases where the transformation is non-invertible. In such cases, the representation is not unique due to the non-trivial null-space of the transformation matrix, and hence the proposed approach employ sparsity as regularizer to obtain the intended representation. Further, to update the dictionary, a fast greedy dictionary selection (GDS) algorithm is pro-

posed, where dictionary is build by selecting the dictionary columns from multiple training candidates. The proposed greedy algorithm is based on sub-modularity property of the DL problem, it is shown that the algorithm achieves at-least a constant fraction of the optimal value of the DL objective function. For problems where the transformation employed in dictionary preconditioning requires to be learned rather than one having analytical or parametric form, a fast exemplar selection (FES) algorithm is proposed. It selects an optimal subset which spans the same space as the whole training data based on the principles of low-rank matrix recovery/approximations. To make the approach scalable to large ensemble of training signals, incremental Cholesky decomposition and block matrix inversion algorithms are employed to speed up the sampling process. In light of the proposed GDL algorithm, experimental results are demonstrated for learning dictionaries in the context of many well known existing matrix factorization problems involving factorisable structure namely double sparse dictionary learning (DSDL), kernel sparse dictionary learning (KSDL), archetypal analysis (AA), functional connectivity analysis in functional magnetic resonance imaging (fMRI) signals. For instance, DSDL problem involve learning a dictionary as a product of an analytical transform and a spare matrix. Similarly, KSDL and AA problems involve learning a dictionary as a product of the training exemplars and a spare matrix.

Finally, the work in this thesis is further extended to use the sparse representation for the inference problem of voiced/nonvoiced (V/NV) detection and signal recovery from compressively sensed speech signals. It attempts to exploit the fact that the inherent glottal activity characteristic of the speech production mechanism is captured in the sparse representation, which can be used for V/NV detection. The estimation of sparse vector is influenced by the dictionary, which is difficult to obtain when only compressed samples of signal are available. To address this, this work proposed a new method which shows that it is indeed possible to build the sparsifying dictionary using only compressive samples. In particular, EMD decompositions of compressive samples are used to form the atoms of the dictionary, and is motivated by the fact that CS samples have envelop similar to the envelop of original samples.

Contents