Evaluating Dictionary Learning for Sparse Representation Algorithms using SMALLbox

Ivan Damnjanovic, Matthew E. P. Davies and Mark D. Plumbley
School of Electronic Engineering and Computer Science
Queen Mary University of London
Mile End Road, London, E1 4NS, UK
Email: {name}.{surname}@eecs.qmul.ac.uk

***Abstract*—SMALLbox is an open source MATLAB toolbox aiming at becoming a testing ground for the exploration of new provably good methods to obtain inherently data-driven sparse models, which are able to cope with large-scale and complicated data.**

I. SMALLbox - Evaluation Framework

The field of sparse representations has gained a huge interest in recent years, in particular in applications such as compressed sensing, image de-noising and source separation. We are witnessing a growing number of sparse representation algorithms that are becoming freely available in the research community [1-2]. This growth raised a necessity for an environment for proper testing and benchmarking. The SPARCO framework [3] partially addresses this problem by providing a large collection of imaging, signal processing, compressed sensing, and geophysics sparse reconstruction problems for testing these algorithms. It also includes a large library of operators that can be used to create new test problems.

Sparse representation approaches find the sparse solution in a given dictionary, but give suboptimal solution in many scenarios in which no suitable model is known. Many algorithms exist that aim to solve the sparse representation dictionary learning problem [4-5]. The main driving force for this work is the lack of a toolbox such as SPARCO for dictionary learning problems. Recognising the need of the community for such a toolbox, we set out to design SMALLbox - a MATLAB toolbox with three main aims:

- to enable an easy way of comparing dictionary learning algorithms,
- to provide a unifying API that will enable interoperability and re-use of already available toolboxes for sparse representation and dictionary learning,
- to aid the reproducible research effort in sparse signal representations and dictionary learning.

To enable re-use of already developed problems from SPARCO, the main interoperability is given through the "Problem" structure which in SMALLbox can be defined either as a sparse representation or dictionary learning problem. In generating a problem, some of the utilities can be used to decode a dataset and prepare a test signal or a training set for dictionary learning. The dictionaries can be either defined or learned using dictionary learning algorithms. In the former case, they can be given as implicit dictionaries, as a combination of the given operators and structures, or explicitly in the form of a dictionary matrix. In the latter case, they are learned from training data. Once the dictionary is set in the problem, the problem is ready to be solved by one of the sparse representation algorithms.

SMALLbox has been designed to enable an easy exchange of information and a comparison of different modules developed through a unified API structure. The structure was made to fulfil two main goals. The first goal is to separate a typical sparse signal processing problem into three meaningful units:

a) problem specification (preparing data for learning the structures, representation and reconstruction),
b) dictionary learning (using a prepared training set to learn the natural structures in the data) and
c) sparse representation (representing the signal with a pre-specified or learned dictionary).

The second goal is to provide a seamless connection between the three types of modules and ease of communication of data between the problem, dictionary learning and sparse representation parts of the structure. To achieve these goals, SMALLbox provides a "glue" structure to allow algorithms from different toolboxes to be used with a common API.

The SMALLbox evaluation framework is implemented as a MATLAB toolbox, which can be downloaded from http://small-project.eu and is in the form of an archive containing the SMALLbox directory structure and necessary MATLAB scripts. To enable easy comparison with the existing state-of-the-art algorithms, installation scripts will download third party toolboxes as required. In addition, the code is well documented with examples giving step-by-step instructions of how to implement new problems or introduce new sparse-representation and dictionary learning algorithms to the toolbox. These examples are built upon the set of test problems already implemented in SMALLbox. Since SMALLbox is an evaluation framework of the EU FET SMALL project, more problems, solvers and dictionary learning techniques that are developed will be included in SMALLbox as the project proceeds.

Acknowledgment

This research is supported by EU FET-Open ProjectFP7-ICT-225913 "SMALL", EPSRC Platform Grant EP/045235/1, and EPSRC "Sustainable Software for Digital Music and Audio Research" Grant EP/H043101/1.

References

[1] D. Donoho, V. Stodden and Y. Tsaig, *Sparselab*, 2007, http://sparselab.stanford.edu/
[2] T. Blumensath and M. E. Davies, *Gradient pursuits*, In IEEE Transactions on Signal Processing, vol. 56, no. 6, pp. 2370-2382, June 2008.
[3] E. v. Berg, M. P. Friedlander, G. Hennenfent, F. Herrmann, R. Saab and O. Yilmaz *Sparco: A testing framework for sparse reconstruction*, In ACM Trans. on Mathematical Software, 354:1-16, February 2009.
[4] R. Rubinstein, M. Zibulevsky and M. Elad, *Double Sparsity: Learning Sparse Dictionaries for Sparse Signal Approximation*, In IEEE Transactions on Signal Processing, Vol. 58, No. 3, Pages 1553-1564, March 2010.
[5] K. Skretting and K. Engang, *Recursive Least Squares Dictionary Learning Algorithm*, In IEEE Transactions on Signal Processing, Vol 58, no4, 2010.