High Dimensional Dictionary Learning and Applications

Jeremias Sulam, Michael Zibulevsky and Michael Elad
Computer Science Department
Technion - Israel Institute of Technology
{jsulam,mzib,elad}@cs.technion.ac.il

***Abstract*—In this work, we show how to efficiently handle bigger dimensions and go beyond the small patches in sparsity-based signal and image processing methods. We build our approach based on a new cropped Wavelet decomposition, which enables a multi-scale analysis with virtually no border effects. Employing this as the base dictionary within a double sparsity model, and to cope with the increase of training data, we present an Online Sparse Dictionary Learning (OSDL) algorithm to train this model effectively, enabling it to handle millions of examples. The resulting large trainable atoms – *trainlets* – not only achieve state of the art performance in dictionary learning when compared to other methods, but it also opens the door to new challenges and problems that remained unattainable until now. In addition to reviewing the capabilities of the OSDL algorithm, we present very recent results on inpainting of large regions of face images, as well as preliminary results on full end-to-end image compression.**

I. Introduction

Sparse representations has shown to be a powerful prior in several inverse problems in image processing. This model assumes that a signal $\mathbf{y} \in \mathbb{R}^n$ can be well approximated by a decomposition of the form $\mathbf{D}\mathbf{x}$, where $\mathbf{D}$, termed dictionary, is a matrix of size $n \times m$ containing signal atoms in its columns, and a sparse vector $\mathbf{x} \in \mathbb{R}^m$. The problem of finding such a sparse vector is termed sparse coding, and is usually formulated in terms of a pursuit algorithm. When combined with the ability to learn the dictionary from real data, and for a specific task, this model has yielded a number of state of the art results [1], [2], [3], [4]. A series of different dictionary learning algorithms have been proposed [5], [6], [1], [7], most of them employing an alternating minimization approach minimizing over the set of sparse representations and $\mathbf{D}$.

However successful, the dictionary learning problem has traditionally been restricted to the domain of small image patches, thus limiting the kind of problem these methods can address. This limitation arises mainly from computational constraints, but also from the fact that the degrees of freedom of the problem – and the amount data required – become unmanageable as the dimension increases. Some works have attempted to provide more efficient dictionary learning algorithms. The work presented in [7], for example, proposed to lower the complexity of using (and learning) the dictionary by suggesting an adaptable but completely separable structure, yielding an algorithm term SEDIL (Separable Dictionary Learning). However interesting, the complete separability constraint is often too restrictive to represent general images of high dimensions, and its batch-learning algorithm is restricted to relatively small training sets.

II. Trainlets

In a very recent work [8], we have presented the Online Sparse Dictionary Learning (OSDL) algorithm, which is able to manage signals of dimensions in the order of the several thousands and beyond. This approach builds on the work of [9], which models the dictionary $\mathbf{D}$ as the product of a fast and efficient *base* dictionary, and an adaptable sparse factor $\mathbf{A}$. This lowers the complexity of both, the degrees of freedom of the problem and the computational cost of applying the dictionary. Formally, the (sparse) dictionary learning problem is formulated as

$$\min_{\mathbf{A},\mathbf{X}} \ \frac{1}{2}||\mathbf{Y} - \mathbf{\Phi}\mathbf{A}\mathbf{X}||_F^2 \quad \text{subject to} \quad \begin{cases} ||\mathbf{x}_i||_0 \leq p & \forall i \\ ||\mathbf{a}_j||_0 = k & \forall j \end{cases}, \qquad (1)$$

where $\mathbf{Y}$ is a matrix having the training examples in its columns, $\mathbf{X}$ contains the corresponding representation vectors, and the matrix $\mathbf{A}$ is column-wise sparse. In particular, we employ a novel Cropped Wavelets dictionary as the operator $\mathbf{\Phi}$. This operator, also introduced in our work, enables an optimal multi-scale decomposition (in a sparse sense) with virtually no border effects. In order to cope with the increase of training data, we suggest a dictionary learning algorithm based on ideas from stochastic optimization [10]. In a nutshell, the algorithm performs sparse coding of a mini-batch of training examples with (Sparse) OMP, and then updates a subset of the dictionary atoms through a variation of the NIHT algorithm [11].

III. Experiments and Applications

The OSDL algorithm enables us to apply dictionary learning on millions of training images of relatively large size, obtaining insightful results. For instance, such a (general purpose) dictionary for 32×32 natural image patches is shown in Figure 1: the OSDL learns very different atoms, as can be seen from the piece-wise-constant ones, to textures at different scales and edge-like atoms. It is interesting to see that Fourier type atoms, as well as Contourlet and Gabor-like atoms, naturally arise from the data. In addition, such a dictionary obtains some flavor of shift (and even rotation) invariance, as similar patterns appear in different locations in different atoms.

One can leverage the representation power of the OSDL algorithm to design global atoms of face images, and then employ these to recover or infer large missing regions in the image. This extreme image restoration problem intrinsically requires some kind of global model, as any other local or patch-based alternative would fail in generating the correct content for this specific context. Figure 2 exemplifies one of our results reported in [12], obtained with an ad-hoc dictionary for face images of size 100×100 (not shown in this paper). Building on a similar model, one can suggest a complete compression algorithm for general face images. Preliminary results for such an approach are presented in Figure 3, outperforming both JPEG and JPEG 2000 compressions algorithms. These, and more, results demonstrate that the OSDL algorithm not only achieves state of the art performance in dictionary learning when compared to other methods (for example, in terms of representation error) but it also opens the door to new challenges and problems that remained unattainable until now.

Acknowledgement

The research leading to these results has received funding from the European Research Council under European Unions Seventh Framework Programme, ERC Grant agreement no. 320649

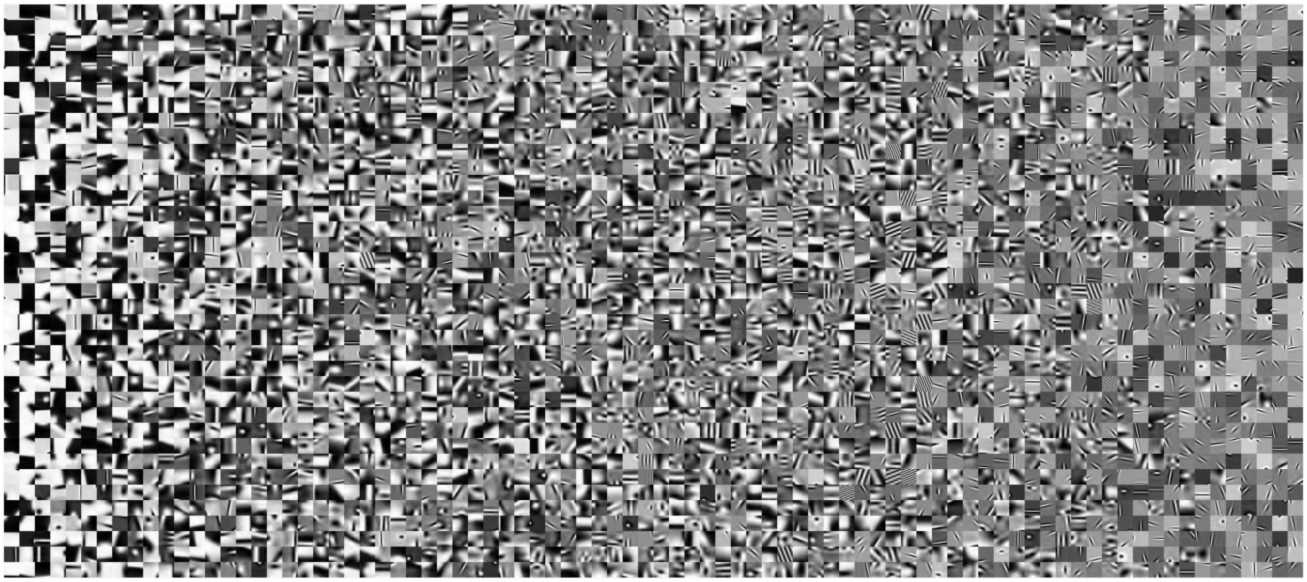

Fig. 1: General purpose dictionary. Subset of the learned atoms on millions on natural of images patches (32×32).

Fig. 2: Inpainting application. From left to right: masked image, patch propagation [13], PCA, SEDIL [7], Trainlets [8], and the original (100×100) image.

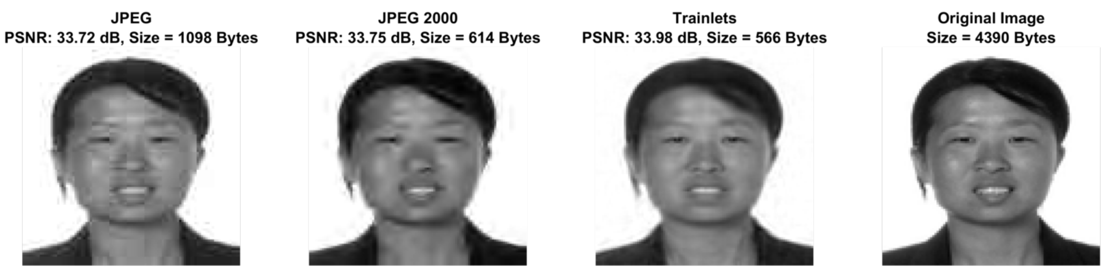

Fig. 3: Face images compression. From left to right, the results of: JPEG, JPEG2000 (not including file-header) and Trainlets, and the original (100×100) image. Test images not including in the training phase.

REFERENCES

[1] J. Mairal, F. Bach, and G. Sapiro, "Non-local Sparse Models for Image Restoration," *IEEE International Conference on Computer Vision.*, vol. 2, pp. 2272–2279, 2009.

[2] Y. Romano, M. Protter, and M. Elad, "Single image interpolation via adaptive nonlocal sparsity-based modeling," *IEEE Trans. on Image Process.*, vol. 23, no. 7, pp. 3085–3098, 2014.

[3] R. Giryes and M. Elad, "Sparsity-based poisson denoising with dictionary learning," *Image Processing, IEEE Transactions on*, vol. 23, no. 12, pp. 5057–5069, 2014.

[4] Z. Jiang, Z. Lin, and L. S. Davis, "Label consistent k-svd: Learning a discriminative dictionary for recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2651–2664, 2013.

[5] K. Engan, S. O. Aase, and J. H. Husoy, "Method of Optimal Directions for Frame Design," in *IEEE Int. Conf. Acoust. Speech, Signal Process.*, pp. 2443–2446, 1999.

[6] M. Aharon, M. Elad, and A. M. Bruckstein, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," *IEEE Trans. on Signal Process.*, vol. 54, no. 11, pp. 4311–4322, 2006.

[7] S. Hawe, M. Seibert, and M. Kleinsteuber, "Separable dictionary learning," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 438–445, 2013.

[8] J. Sulam, B. Ophir, M. Zibulevsky, and M. Elad, "Trainlets: Dictionary learning in high dimensions," *IEEE Transactions on Signal Processing*, vol. 64, no. 12, pp. 3180–3193, 2016.

[9] R. Rubinstein, M. Zibulevsky, and M. Elad, "Double Sparsity : Learning Sparse Dictionaries for Sparse Signal Approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1553–1564, 2010.

[10] L. Bottou, "Online algorithms and stochastic approximations," in *Online Learning and Neural Networks*, Cambridge University Press, 1998. revised, Oct 2012.

[11] T. Blumensath and M. E. Davies, "Normalized iterative hard thresholding: Guaranteed stability and performance," *IEEE Journal on Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 298–309, 2010.

[12] J. Sulam and M. Elad, "Large inpainting of face images with trainlets," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1839–1843, 2016.

[13] Z. Xu and J. Sun, "Image inpainting by patch propagation using patch sparsity," *Image Processing, IEEE Transactions on*, vol. 19, no. 5, pp. 1153–1165, 2010.