Hierarchized block wise image approximation by greedy pursuit strategies

Laura Rebollo-Neira, Ryszard Macioł and Shabnam Bibi
Mathematics Department, Aston University, Birmingham, B4 7ET, UK

Abstract—**An approach for effective implementation of greedy selection methodologies, to approximate an image partitioned into blocks, is proposed. The method is specially designed for approximating partitions on a transformed image. It evolves by selecting, at each iteration step, i) the elements for approximating each of the blocks partitioning the image and ii) the hierarchized sequence in which the blocks are approximated to reach the required global condition on sparsity.**

I. Introduction

The representation of an image by a piece of data, of lower dimensionality than the pixel/intensity values, is called a sparse representation of a compressible image. Approximations using a redundant set, called a *dictionary*, may attain high sparsity, thereby benefiting applications that range from denoising to compression [1]–[9]. Those techniques for approximation which evolve by selection of dictionary elements, called *atoms*, are refereed to as greedy strategies. The motivation of this Communication is to consider an effective implementation of the Orthogonal Matching Pursuit (OMP) strategy [10] to approximate an image in the wavelet domain. For this to be computationally effective, with respect to speed and storage demands, the transformed image should be partitioned into small blocks.

Approximations in the wavelet domain have been shown useful for image compression [5]–[7]. Indeed, as will be illustrated here, the approximation of images which are compressible in the wavelet domain results significantly sparser if carried out in such a domain. In addition, while some other artifacts can be caused, approximations by blocking in the wavelet domain avoid visually unpleasant blocking artifacts in the intensity image.

The letter is organized as follows: Sec. II discusses the need for considering a dedicated version of the OMP strategy to operate on a transformed image. Sec. III proposes a particular version, which is termed hierarchized block wise OMP. The advantage of the proposed approach is illustrated by numerical tests in Sec. IV. The conclusions are summarized in Sec. V.

II. The need for dedicated pursuit strategies to operate on partitions

Approximating an image partitioned into blocks implies having to consider some 'communication' between the blocks. The need for communication appears when trying to decide up to what error each block is to be approximated. As the following discussion suggests, an appropriate stopping criterion for recursive approximation of blocks in the wavelet domain needs to be a *global measure*.

Firstly let us introduce the notational convention: $\mathbb{R}$ represents the sets of real numbers. Boldface fonts are used to indicate Euclidean vectors or matrices, whilst standard mathematical fonts indicate the components, e.g., $\mathbf{c} \in \mathbb{R}^N$ is a vector of components $c(i)$, $i = 1, \ldots, N$ and $\mathbf{I} \in \mathbb{R}^{N_x \times N_y}$ a matrix of elements $I(i,j)$, $i = 1, \ldots, N_x$, $j = 1, \ldots, N_y$. In particular $\mathbf{I} \in \mathbb{R}^{N_x \times N_y}$ will indicate an intensity image of $N_x \times N_y$ pixels and $\tilde{\mathbf{I}} \in \mathbb{R}^{N_x \times N_y}$ its corresponding wavelet transform. For square matrices i.e., when $N_x = N_y$, to shorten notation the range of indices is indicated as $i,j = 1, \ldots, N_x$.

Consider, without loss of generality, a real, orthogonal, and separable wavelet transform, which transforms an intensity image $\mathbf{I} \in \mathbb{R}^{N_x \times N_y}$ into the transformed image $\tilde{\mathbf{I}} \in \mathbb{R}^{N_x \times N_y}$. Thus, each of the points $\tilde{I}(n,m)$, $n = 1, \ldots, N_x$, $m = 1, \ldots, N_y$ is a Frobenius inner product $\tilde{I}(n,m) = \langle \Psi_{u_n} \otimes \Psi_{v_m}, \mathbf{I} \rangle_F$, where the operation $\otimes$ indicates the tensor product. The subscript u_n in Ψ_{u_n} is used to identify an ordered pair consisting of the scale and translation parameters of the corresponding mother wavelet Ψ. Accordingly, for $n = 1, \ldots, N_x$, $m = 1, \ldots, N_y$,

$$\tilde{I}(n,m) = \sum_{i,j=1}^{N_x, N_y} \Psi_{u_n}(i) I(i,j) \Psi_{v_m}(j). \quad (1)$$

We shall assume for simplicity a uniform partition and consider that a transformed image, $\tilde{\mathbf{I}}$, is the composition of Q identical and disjoint blocks $\tilde{\mathbf{I}}_q$, $q = 1, \ldots, Q$. Hence, $\tilde{\mathbf{I}} = \cup_{q=1}^{Q} \tilde{\mathbf{I}}_q$, where every $\tilde{\mathbf{I}}_q$ is a 2D block of size $N_b \times N_b$.

It is clear from (1) that points of the transformed image corresponding to a particular block of size $N_b \times N_b$ contain some global information about the intensity image. This suggests that a

suitable stopping criterion for the approximation of a partition in the wavelet domain needs to be a global measure. Consequently, even if the approximation of each block, as such, is carried out independently of the others, a greedy selection strategy implemented by approximating blocks of the transformed image should aim at selecting i) the elements for approximating each of the blocks in the partition and ii) the hierarchized sequence in which the blocks should be approximated to reach the global condition required by the algorithm. The method introduced in the next section, which we term hierarchized block wise OMP (HBW-OMP) implements the selection of i) and ii) simultaneously.

III. HIERARCHIZED BLOCK WISE OMP

Given a redundant 2D dictionary $\mathcal{D}$ of M atoms, $\mathcal{D} = \{\mathbf{D}_n \in \mathbb{R}^{N_b \times N_b}\}_{n=1}^M$, each of the Q blocks $\tilde{\mathbf{I}}_q \in \mathbb{R}^{N_b \times N_b}$ in the partition of the transformed image is approximated by an atomic decomposition $\tilde{I}^{k_q,q}(i,j)$, $i,j = 1, \ldots, N_b$ of the form

$$\tilde{I}^{k_q,q}(i,j) = \sum_{n=1}^{k_q} c^{k_q,q}(n) D_{\ell_n^q}(i,j), \; q = 1, \ldots, Q.$$

For each q, the atoms $\mathbf{D}_{\ell_n^q}$, $n = 1, \ldots, k_q$ are selected from the dictionary $\mathcal{D}$ as follows: On setting $\mathbf{R}^{0,q} = \tilde{\mathbf{I}}_q$, $q = 1, \ldots, Q$ at each iteration the algorithm selects the atom $\mathbf{D}_{\ell_{k_q+1}^q}$ that maximizes the absolute value of the Frobenius inner products $\langle \mathbf{D}_n, \mathbf{R}^{k_q,q} \rangle_{\mathrm{F}}$, $n = 1, \ldots, M$, $q = 1, \ldots, Q$, i.e.,

$$\ell_{k_q+1}^q = \underset{\substack{n=1,\ldots,M \\ q=1,\ldots,Q}}{\arg\max} |\langle \mathbf{D}_n, \mathbf{R}^{k_q,q} \rangle_F|$$

$$\text{with} \tag{2}$$

$$\mathbf{R}^{k_q,q} = \tilde{\mathbf{I}}_q - \sum_{n=1}^{k_q} c^{k_q,q}(n) \mathbf{D}_{\ell_n^q}.$$

For each q and k_q the coefficients $c^{k_q,q}(n)$, $n = 1, \ldots, k_q$ in (2) are such that $\|\mathbf{R}^{k_q,q}\|_{\mathrm{F}}$ is minimum, where $\| \cdot \|_{\mathrm{F}}$ is the norm defined through the Frobenius inner product. This is ensured by requesting that $\mathbf{R}^{k_q,q} = \tilde{\mathbf{I}}_q - \hat{P}_{\mathbb{V}_{k_q}^q} \tilde{\mathbf{I}}_q$, where $\hat{P}_{\mathbb{V}_{k_q}^q}$ is the orthogonal projection operator onto $\mathbb{V}_{k_q}^q = \text{span}\{\mathbf{D}_{\ell_n^q}\}_{n=1}^{k_q}$. The implementation discussed in [11] provides us with the representation of $\hat{P}_{\mathbb{V}_k^q} \tilde{\mathbf{I}}_q$ as given by,

$$\hat{P}_{\mathbb{V}_k^q} \tilde{\mathbf{I}}_q = \sum_{n=1}^{k_q} \mathbf{D}_{\ell_n^q} \langle \mathbf{B}_n^{k,q}, \tilde{\mathbf{I}}_q \rangle_F = \sum_{n=1}^{k_q} c^{k,q}(n) \mathbf{D}_{\ell_n^q}.$$

For q fixed, the matrices $\mathbf{B}_n^{k_q}$, $n = 1, \ldots, k_q$ are biorthogonal to the selected atoms $\mathbf{D}_{\ell_n^q}$, $n = 1, \ldots, k_q$ and span the identical subspace $\mathbb{V}_{k_q}^q$. These matrices can be effectively calculated by recursive biorthogonalization and Gram Schmidt orthogonalization with one re-orthogonalization step.

The coefficients in (2) are obtained from the inner products $c^{k_q,q}(n) = \langle \mathbf{B}_n^{k_q,q}, \tilde{\mathbf{I}}_q \rangle_{\mathrm{F}}$, $n = 1, \ldots, k_q$.

For a given number K, the algorithm iterates until the condition $\sum_{q=1}^Q k_q = K$, is met. In other words, the algorithm stops when the maximum number of *total atoms* required for the image approximation is reached.

The difference between OMP and the hierarchized block wise (HBW) version discussed here lies in a) the sequence in which the blocks are *partially* approximated at each step and b) the stopping criterion. Standard OMP would be applied independently to each block up to a given error, which is independent of the approximation of the other blocks. As condition (2) states, HBW-OMP adds a hierarchized selection of the blocks to be approximated in each iteration. Thus, at each step, the maximum in (2) changes only for the selected block. This implies that, by storing the maximum for each block, the selection of blocks introduces the overhead of finding the maximum element of an array of size Q. As far as storage is concerned, HBW-OMP has to store all the stepwise outputs of the algorithm, which includes matrices $\mathbf{B}_n^{k_q}$, $n = 1, \ldots, k_q$, for each of the Q blocks in the partition. While the approximation is carried out in a block wise manner, the relation introduced by the condition $\sum_{q=1}^Q k_q = K$ inhibits the complete approximation of each block at once. Other pursuit strategies can also be adapted to this selection process. In particular, the HBW implementation of the Matching Pursuit (MP) method [12] differs from that of OMP in that, for each q and k_q, the coefficients $c^{k_q,q}(n)$, $n = 1, \ldots, k_q$ in (2) are calculated simply as $c^{k_q,q}(n) = \langle \mathbf{D}_{\ell_n^q}, \mathbf{R}^{k_q-1,q} \rangle_F$.

It is appropriate to highlight also the difference between the proposed approach and the Block OMP/MP (BOMP/BMP) approach introduced in [13]. While BOMP/BMP selects *blocks of atoms* at each iteration step, our proposal keeps selecting the atoms as in OMP/MP , i.e., one by one.

IV. NUMERICAL TESTS

The dictionary used for the tests is a separable mixed dictionary consisting of two components, $\mathcal{D}_1$ and $\mathcal{D}_2$. The component $\mathcal{D}_1$ is a Redundant Discrete Cosine (RDC) dictionary given by:

$$\mathcal{D}_1 = \{w_i \cos(\frac{\pi(2j-1)(i-1)}{2M}), \; j = 1, \ldots, N\},$$

with w_i, $i = 1, \ldots, M$ normalization factors. The number N is set equal to the number of pixels in one of the block sides, and $M = 2N$ to have redundancy two. $\mathcal{D}_2$ is the standard Euclidean basis, also called the Dirac basis (DB), i.e.

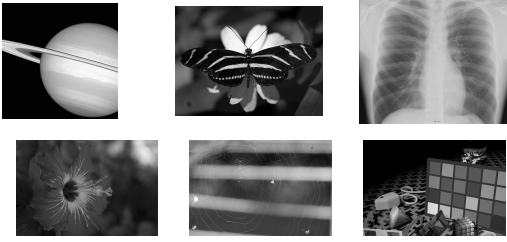$$\mathcal{D}_2 = \{e_i(j) = \delta_{i,j}, \; j = 1, \ldots, N\}_{i=1}^N.$$

Fig. 1. First row: Planet, Butterfly and Chest X-Ray images of sizes 512×512, 448×600, and 592×728 pixels, respectively. Second row: Flower, Spider Web and Artificial high resolution images from [18], all of size 1512×2264 pixels.

I	MP	HBW	OMP	HBW	WT	DCT
P	39.1	53.0	43.9	62.9	40.7	27.1
B	10.6	12.0	12.7	14.4	9.6	10.7
C	32.5	53.2	35.0	60.7	34.2	20.7
F	48.6	156.6	50.5	181.5	119.5	40.0
S	41.1	87.9	43.3	99.2	70.0	33.5
A	25.0	30.1	28.4	35.0	24.9	19.4

TABLE II

SAME DESCRIPTION AS IN TABLE 1 BUT IN THIS CASE MP, OMP AND THE CORRESPONDING HBW VERSIONS ARE IMPLEMENTED IN THE WAVELET DOMAIN. THE RESULTS OF THE WT AND DCT ARE REPEATED HERE TO FACILITATE THE COMPARISON.

The required 2D dictionary is the tensor product $\mathcal{D} \otimes \mathcal{D}$, where $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$. This small dictionary, called RDCDB in [8], is chosen only for simplicity. Extending the RDCDB dictionary by adding wavelet atoms, for instance [14], or using other dictionaries such as those arising by learning processes [15]–[17], would result in an increment of sparsity with all the approaches. However the purpose of these numerical tests is only to illustrate the advantage of implementing matching pursuit strategies on partitions, in the proposed HBW manner. The set of six grey intensity levels images used for illustrating the proposed HBW strategy are shown in Fig 1. The images in the first row: Planet, Butterfly and Chest X-Ray are of sizes 512×512, 448×600 and 592×728 pixels. The images in the second row: Flower, Spider Web and Artificial are high resolution images taken from [18]. The Spider Web and Artificial are both only fractions of the much larger images in [18], which we have cropped to the same size as the Flower, i.e., 1512×2264 pixels.

All the images of Fig. 1 are approximated in both, the intensity and wavelet domains by partitions of 8×8 pixels. The transformed image is obtained using the CDF97 WT. The sparsity of an approximation is measured by the Sparsity Ratio (SR) defined as $\text{SR} = \frac{N_x N_y}{K}$, with $N_x N_y$ the total number of pixels and K the number of nonzero coefficients to approximate the whole image. The

I	MP	HBW	OMP	HBW	WT	DCT
P	28.1	29.0	30.4	31.0	40.7	27.1
B	10.6	11.2	12.3	12.7	9.6	10.7
C	22.6	23.2	23.6	24.4	34.2	20.7
F	41.0	42.8	42.6	43.9	119.5	40.0
S	34.8	35.7	36.5	37.0	70.0	33.5
A	26.4	27.5	29.7	30.7	24.9	19.4

TABLE I

SR RESULTING FROM APPROXIMATING (IN THE INTENSITY DOMAIN) THE SIX IMAGES OF FIG. 1, UP TO PSNR=45.0DB, BY THE METHODS: MP, OMP, THEIR CORRESPONDING HBW VERSIONS, WT AND DCT.

approximation of all the images is carried out to

achieve the required PSNR of 45.0dB using the above defined RDCDB dictionary and the greedy strategies OMP, MP and their corresponding HBW versions. The SRs arising when these techniques are implemented in the intensity domain are displayed in the first four columns of Table 1. The rows, from top to ground, correspond to the Planet (P), Butterfly (B), Chest (C), Flower (F), Spider Web (S) and Artificial (A) images (I). For comparison purposes, the results produced by conventional WT and DCT approaches are displayed in the last two columns of the table. The WT results are obtained by applying the CDF97 WT, on the whole image, and reducing coefficients by iterative thresholding until the PSNR of 45.0dB is reached. The results from the orthogonal DCT are obtained, from a partition of 8×8 pixels, also by thresholding of coefficients.

From Table 1 it is clear that: a) Except for the Butterfly and Artificial images, in the intensity domain the results produced by a conventional WT approach significantly over perform the other approaches. b) Whilst the block independent OMP and MP methods produce smaller SRs than their corresponding HBW versions, the difference is not very significant. However, as can be observed in Table 2, all this is reversed if the greedy techniques are implemented with the RDCDB dictionary but in the wavelet domain. In this domain: a) for all the images the HBW versions of MP and OMP, with the simple RDCDB dictionary, significantly over-perform conventional WT and DCT approaches. b) Except for the Butterfly and Artificial images the difference in the SR produced by the HBW approaches, with respect to the block independent versions, is <u>massive</u>.

Putting aside the Butterfly and Artificial images, the common feature of the other images in Fig. 1 is the very significant difference in the SR produced by a conventional WT approximation, with respect to the conventional DCT. This is an indication that the representation of the images is particularity sparse in the wavelet domain. It is clear from Table 2 that, the sparser the approximation by the WT is,