MULTI-STAGE OMP Sparse Coding Using Local Matching Pursuit Atoms Selection

[1]Mahdi Marsousi, [2]Kaveh Abhari, [4]Paul Babyn, [2,3]Javad Alirezaie, Senior Member, IEEE

[1]Department of Electrical and Computer Engineering, University of Toronto, Canada
[2]Department of Electrical and Computer Engineering, Ryerson University, Canada
[3]Department of Systems Design Engineering, University of Waterloo, Canada
[4]Department of Medical Imaging, University of Saskatchewan and Saskatoon Health Region, Canada

ABSTRACT

A new multi-stage approach based on component extraction is proposed to more efficiently address the sparse representation problem. In each stage a pre-set number of coefficients are chosen for reconstructing each signal component. A global search is performed to extract a lower dimensional sub-dictionary consisting of a sorted set of candidate atoms to represent the signal component, corresponding to the stage. The best representing atoms are then selected from the sub dictionary using the Matching Pursuit (MP) method. Afterwards, the sparse coefficients are updated in the same manner in which the Orthogonal Matching Pursuit (OMP) operates. The proposed method is more efficient that the conventional OMP methods. To evaluate the performance of the proposed method, it is compared to OMP and Stagewise OMP (StOMP), which are conceptually the most similar to the proposed approach. The results illustrate the proposed method is more time efficient than OMP and more robust and sparser than the StOMP.

Index Terms— multi-stage, sparse representation, Orthogonal Matching Pursuit (OMP), StOMP, MP

1. INTRODUCTION

Sparse coding theorem suggests that signals can be efficiently represented as a weighted linear combination of a set of prototype elementary vectors. These basic vectors are called atoms and form a dictionary. Sparse representation does not have a unique solution and is proven to be an NP hard [1] problem. Several methods are proposed to find the sparsest solution. These methods can be categorized into four classes, namely: 1) Greedy pursuit 2) Convex and non-convex optimization 3) Probabilistic modeling 4) Exhaustive search. In this paper we propose to find the sparsest solution based on Greedy pursuit algorithms. Greedy algorithms generally start by approximating a rough estimate of the signal and improving the approximation recursively until a predefined precision is met. The famous greedy sparse coding methods include MP [2], OMP [3] and its alternatives such as Batch OMP, Stagewise OMP (StOMP) [4] and Regularized OMP (ROMP) [5].

Given an over-complete dictionary, MP decomposes a signal by finding the locally optimum solution for each coefficient calculation. MP isolates the atom that best resembles the signal by means of its projection on the dictionary (inner product), and then subtracts the contribution of the specific atom. The process is repeated until a predefined criterion is met. OMP is an extension of MP with the main difference that the coefficients are the orthogonal projection of the signal on the dictionary and the coefficients are accordingly updated at each iteration. Even though MP is more feasible to implement, OMP results in more accurate sparse approximations. In these methods a global search on the dictionary performed for each coefficient and the coefficients are calculated and updated one at a time. This fact leads to a relatively slow process. In order to overcome this issue, new approaches based on OMP are proposed. In these methods the solution is achieved by adding multiple atoms at a time while OMP solution is built by one-by-one addition of coefficients. Among these methods, StOMP is one the most common techniques and is the most similar approach to the proposed method.

StOMP calculates the projection of the signal on the dictionary then in each stage selects all potential atoms above a preset threshold value and uses a least squares method to find an approximation in each stage. This is then repeated with a residue vector. This results in a faster convergence of the solution. However, due to use of hard thresholding, the sparsity is not controlled and results may vary for different values of threshold.

In this paper, we propose a new multi-stage sparse coding approach (MS-OMP) by decomposing the signal into multiple components first and solving for each component in each stage to improve the efficiently while maintaining the sparsity criterion of the solution. In each stage a pre-set number of coefficients are chosen for reconstructing each signal component. A global search is performed to extract a sorted set of candidate atoms to form a lower dimension sub dictionary, which is utilized in representing the signal component, corresponding to the stage. The best

representing atoms are then selected from the sub dictionary using the MP method. The selected coefficients are then updated through OMP procedure.

The paper is organized as follows; in section 2, the mathematical approach and methodology is explained in detail. The obtained results of applying the proposed method on signals and natural images, as well as a comparison of methods are presented in section 3. Finally, section 4 is the conclusion and our future prospect of the approach.

2. METHODOLOGY

2.1. Component-Based definition

For an input patch, $\overline{P} \in R^N$, there exists a linear combination of functions (components), $\{\overline{f}_k \in R^N\}$,

$$\overline{P} = \sum_{k=1}^{K} \overline{f}_k \tag{1}$$

where a set of dictionary atoms can be found in which the inner product of k^{th} and $(k+1)^{th}$ components satisfies the following condition,

$$\begin{aligned} &\left| \left\langle \overline{f}_k \cdot d_{i_a} \right\rangle - \left\langle \overline{f}_k \cdot d_{i_b} \right\rangle \right| < t_1, \; i_a, i_b \in U_k \\ &\textit{and,} \\ &\left\langle \overline{f}_k \cdot d_i \right\rangle - \left\langle \overline{f}_{k+1} \cdot d_j \right\rangle > t_2, \quad j \in U_{k+1} \text{ and } i \in U_k \end{aligned} \tag{2}$$

where t_1 and t_2 are two values which satisfying $t_2 >> t_1$. Notation $\langle . \rangle$ stands for the inner product of two vectors and U_k is a set of dictionary atom indexes which satisfies the above condition for the k^{th} component, $\overline{f}_k$. With this definition, the inner product of the input patch and the dictionary atoms result in maximum values corresponding to $\overline{f}_1$ which are members of U_1. Maximum values related to the second component, $\overline{f}_2$, which are members of U_2 are achieved using the inner product of $res_1 = \overline{P} - \overline{f}_1$ and dictionary atoms. Other $\{U_k, k > 2\}$ are obtained using the inner product of res_{k-1} and atoms,

$$res_{k-1} = \overline{P} - \sum_{n=1}^{k-1} \overline{f}_n \tag{3}$$

Having $U = \{U_k, k = 1, ..., K'\}$ where $K' < K$, an approximation of $\overline{P}$, $\tilde{P} = \sum_{k=1}^{K'} \overline{f}_n$, is attained using the pseudo-inverse equation as,

$$\begin{aligned} &\tilde{\alpha} = \left(D_U^T D_U \right)^{-1} D_U^T P \\ &\tilde{P} = D_U \tilde{\alpha} \end{aligned} \tag{4}$$

2.2. Multi-Stage Component Extraction

A multi-stage method where in the k^{th} stage, k components are extracted out of the input patch is developed. At the stage S, the inner product of dictionary atoms and res_{S-1}, assuming $res_0 = \overline{P}$, is computed to find candidate atoms entering the U_S to update $U = \{U_n, n = 1, ..., S\}$ at the S^{th} stage. The L highest values of the inner product correspond to the candidate atoms. The candidate atoms are obtained and sorted using a fast sorting method described in section 2.3. These candidate atoms have higher matching with the k^{th} component and they form a local dictionary, D_L. The MP method [2] seeks among the D_L to find a combination of fix number, M, of atoms ($M \le L$) to efficiently represent res_{S-1} in each stage to approximate the corresponding component. After selecting M representing atoms of S^{th} stage, $U = \{U_n, n = 1, ..., S\}$ is updated and $\tilde{\alpha}$ is calculated using (4). This step resembles the update step of OMP while finding the representing atoms are handled through MP method atom finding strategy. It should be noted that for $M = 1$, the result of this method is exactly equal to the OMP method. After a specific limit, increasing the value of M is not effective to reduce the component representing error and hence, a good selection of M is a tradeoff between accuracy and computational cost. Fig 1 shows the block diagram of the proposed sparse representation method. The pseudo-algorithm of the proposed method is presented in Table I.

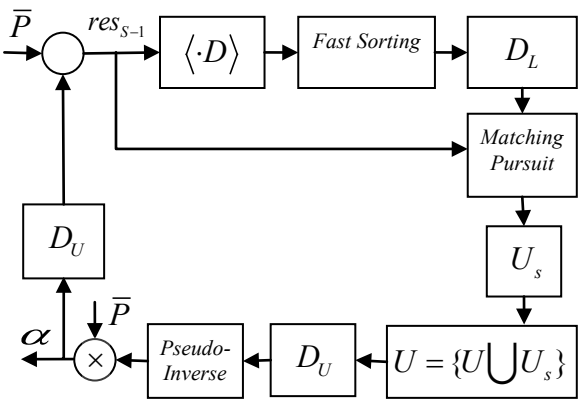

Fig. 1. Block Diagram of the proposed method

2.3. Fast sorting method

This method is developed to efficiently find L highest values among all inner products of dictionary atoms and the res_{S-1} at the S^{th} stage. First, a set of indexes and values $T \in R^{2\times L}$ is initialized with all values equal to zero where the first row refers to maximum values and the second row refers to corresponding atom indexes. The inner product values are checked one by one and if the inner product value corresponding to atom number m satisfies $v^m > T(1,L)$, this atom enters into T and its order within the set is determined using the bubble sort method [6]. The column L containing the lowest value, is eliminated at the end of sorting procedure.

3. RESULTS

As described previously, our method is inspired by the OMP method intended to increase the efficiency of the sparse representation for large size dictionaries. The StOMP method has basically a close definition to our proposed method. However, there exist some essential differences between the two techniques. The threshold value of 0.6 was used for the StOMP method. In order to make a realistic computational time comparison between our proposed method (MS-OMP), the OMP and the StOMP, we have implemented these algorithms using the Microsoft C# .Net. A fixed DCT dictionary consisting of 841 atoms is used to represent images. To provide a comparable analysis to other existing methods, we used standard images including: Lena, Barbara, Baboon, Goldhill, Cameraman, Peppers. The applied dictionary is PI-DCT introduced in [7] which consists of 841 atoms. The size of extracted patches is 8×8 pixels. The sensitivity of the proposed method to two parameters, including M and L, is assessed using the standard images and the result is presented in Fig 2. Accordingly, the selected values are $M=3$ and $L=10$. Our evaluation is mainly focused on two factors: The computational time (CT) and peak signal to noise of the reconstructed image (PSNR).

Table I. Pseudo-Algorithm of the proposed method

Init: $\overline{P}$, $res_0 = \overline{P}$, $\overline{\alpha} \to \overline{0}$, D_U

for S=1:N

$proj = \langle res_{S-1}, D \rangle$

$ind = fastSorting(proj)$

$D_L = \{d_{ind(1)}, ..., d_{ind(L)}\}$

$res2 = res_{S-1}$

$U_s \to empty$

for n=1: M

$m2 = \langle res, D_L \rangle$

$ind2 = \arg\max(m2)$

$U_s = \{U_s \bigcup ind(ind2)\}$

$d' \to D_L(ind2)$

$res2 = res2 - \dfrac{\langle res2, d' \rangle}{\langle d', d' \rangle} d'$

end

$U = U \bigcup U_s$

$\overline{\alpha} = (D_U^T D_U)^{-1} D_U \overline{P}$

$res_{S-1} = \overline{P} - D_U \overline{\alpha}$

end

We have applied the proposed method on the Lena, Cameraman and Peppers images. The average of the results in terms of the accuracy and time efficiency are summarized in Table II.

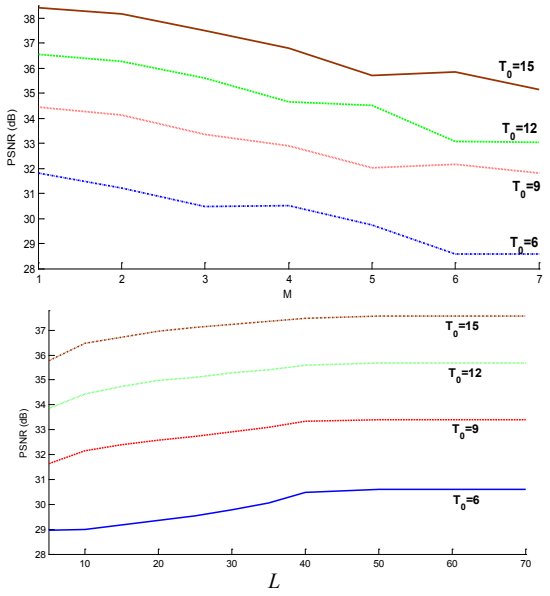

Fig 2. The upper figure shows the sensitivity of the proposed method to M and the bottom figure shows the sensitivity to L.

Table II
Compare average accuracy (PSNR in dB) and average computation time (Sec) for different number of non-zero coefficients (T_0)

T_0	OMP		MS-OMP		StOMP	
	PSNR	CT	PSNR	CT	PSNR	CT
6	41.86	2.56	39.79	2.08	38.39	0.62
9	42.70	4.34	41.49	3.17	38.40	0.82
12	43.31	6.58	42.39	4.62	38.40	0.93
15	43.72	9.44	43.09	6.10	38.41	1.12
18	44.01	13.06	43.59	8.02	40.54	1.48
21	44.15	17.31	43.94	10.64	40.97	1.95
24	44.20	23.63	44.16	12.99	41.08	2.06

According to the result presented in Table II, the proposed method provides higher accuracy than the StOMP while its time of computation is better than the conventional OMP method. The St-OMP method needs much more non-zero coefficients to represent the input signal. Fig 3 demonstrates that the efficiency of using the MS-OMP method increases by increasing the dictionary size in comparison with the conventional OMP approach. Result presented in Fig 3 are obtained using the overcomplete PI-DCT dictionary with sizes from 1000 to 10000. Although the StOMP approach shows faster performance than the MS-OMP, it suffers from a low sparsity level in contrast with the OMP and the MS-OMP methods.

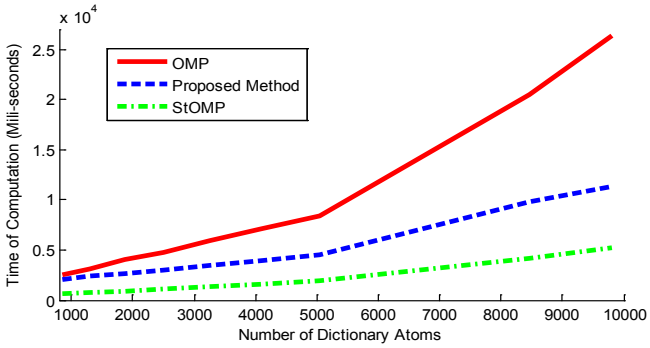

Fig 3. Comparing Time of computations of OMP, MS-OMP and St-OMP methods versus the number of atoms.

It can be observed from the results that the MS-OMP algorithm reduces the computational time as well as maintaining high PSNR as compared to OMP. Examples of qualitative results on the Lena image are shown in Fig 4.

4. CONCLUSION

The proposed method offers a new approach for the sparse representation. It engages a multi-stage approach in which signal components are extracted from the residual of the signal. Our proposed method reduces the number of global searches and substitutes them with local searches on a lower size dictionary. Therefore, the proposed method in comparison with the OMP method needs less computational cost to find the sparse representation considering the same number of non-zero coefficients. According to the definition of the StOMP method, it applies a Hard-Threshold to select entering atoms to the sparse representation and it implies a strong dependency to the selected threshold value. The higher threshold results in lower performance while the lower threshold let a lot of atoms to involve in the sparse representation which results in lower sparsity. In other words, it is difficult to achieve a desired sparsity. In comparison, our proposed method employs a sorting algorithm followed by a matching pursuit search to maximize the efficiency while controlling the sparsity of signal representation. As a prospective enhancement, the adaptive selection of the parameter M according to the representation error can provide more efficiency and accuracy for the proposed method.

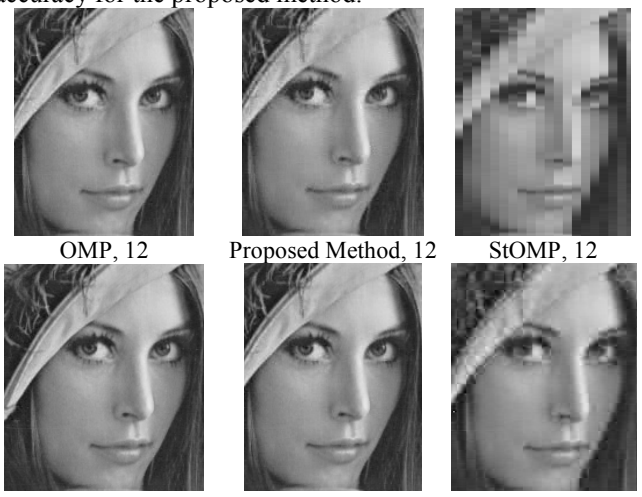

Fig 4. Demonstration of the accuracy of the OMP, proposed method, the StOMP. Top row and bottom row images are the result of representation using 12 non-zero coefficients and 18 non-zero coefficients, respectively.

REFERENCES

[1] G. Davis, S. Mallat, and M. Avellaneda, "Greedy adaptive approximation," *J. Constr. Approx.*, vol. 13, pp. 57–98, 1997.

[2] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing,* vol. 41**,** pp. 3397-3415,1993.

[3] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit." *IEEE Trans. Info. Theory,* vol. 53, pp. 4655–4666, 2007.

[4] D. L. Donoho, Y. Tsaig, I. Drori, J. L. Starck, "Sparse Solution of Underdetermined Systems of Linear Equations by Stagewise Orthogonal Matching Pursuit," *IEEE Transactions on Information Theory,* vol.58, no.2, pp.1094-1121, 2012.

[5] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via

regularized orthogonal matching pursuit," *IEEE Journal of Selected Topics in Signal Processing,* vol. 4, pp. 310-316, April 2010.

[6] A. Owen, "Bubble sort: an archaeological algorithmic analysis," Proceedings of the 34th SIGCSE technical symposium on Computer science education, February 19-23, 2003

[7] K. Abhari, M. Marsousi, P. Babyn, and J. Alirezaie, "Medical Image Denoising Using Low Pass Filtering in Sparse Domain," in *34th Annual International IEEE EMBS Conference*, pp 114-117, 2012.