Sparse Dictionaries Learning for Fast Example-Based Super-Resolution

Qianying Zhang^{1*} and Xiaozhen Xie²

¹Shenzhen Tourism College of Jinan University, Shenzhen 518053, China ²College of Science, Northwest A&F University, Yangling 712100, China zhang_qy@sz.jnu.edu.cn

Abstract

This paper presents a novel approach for the fast example-based super-resolution based upon sparse dictionaries. The proposed method uses a sparse model to learn dictionaries, where the dictionary atoms have sparse representations over a base dictionary. And the batch orthogonal matching pursuit algorithm is used for the sparse-coding operations. The effectiveness of sparse dictionaries is manifested by both well reconstructed quality and speed improvement. We also employ an anchored neighborhood regression to achieve the fast super-resolution. The experimental comparisons with current state-of-the-art algorithms demonstrate the effectiveness of the proposed approach.

Keywords: Super-resolution, Example-based, Sparse representation, Dictionary learning

1. Introduction

Image super-resolution (SR) aims at reconstructing a high-resolution (HR) image from one or a series of low-resolution (LR) observing images. This technique is essential in many practical applications, *e.g.* computer vision, medical image processing and remote sensing. Existing image SR researches can be divided into three main approaches: interpolation-based methods [1, 2], reconstruction-based methods [3, 4] and machine learning (ML) based methods [5, 6]. The latter has been shown to achieve preferable results by using ML techniques. Here we mainly focus on the recent and popular example-based learning SR methods.

Example-based methods use a patch-based approach to learn the relationship between HR and LR image patches, and then obtain the HR estimation. These methods are mostly based on the work of Freeman *et al.* [6]. [6] proposes to construct the HR prediction of the LR image through a Markov random field (MRF). The fundamental example-based methods are based upon neighbor embedding (NE). Cheng *et al.* [7] assume that HR and LR image patch spaces have the similar manifold, and adopt locally linear embedding [8] to estimate HR image patches as a linear combination of neighbors. To reduce the computational complexity, [9] extends this approach by using a Least Squares approximation of LR patches with non-negativity weights. Zhang *et al.* [10] propose a sparse neighbor selection method by clustering the histograms of oriented gradients of LR image patches. The method in [11] exploits the correlation between HR and LR patch features via canonical correlation analysis.

Another class of the example-based method is based on sparse coding [12-14], which attempts to find a sparse representation with respect to an over-complete dictionary for image patches. Yang *et al.* [12] first propose to solve image SR based on this idea. In [12], two HR and LR dictionaries are learned to guarantee the HR and LR patches have the same sparse representation over their corresponding dictionaries. Zeyde *et al.* [13] modify the dictionary learning method, and obtain improvement in both computational

ISSN: 2005-4254 IJSIP Copyright © 2017 SERSC complexity and simulation results. Dong *et al.* [14] propose to learn various sub-dictionaries, and select the sparse domain for each input LR image patch adaptively. Compared to NE-based methods, these methods no longer limit the number of dictionary atoms needed, and adopt sparse coding algorithms to solve the representation coefficients.

The above two classes of example-based methods both rely on manifold similarity between HR and LR patch spaces. Being different from them, the third class method aims at learning a regression function to construct the relationship between HR and LR image patches. Kim *et al.* [15] propose to learn a map from input LR images to target HR images via kernel ridge regression. To achieve fast image SR, the method presented in [16] learns a robust first-order regression model for image SR based on in-place examples. In [17], instead, an anchored neighborhood regression (ANR) model is presented by combining the ideas of NE and sparse coding methods.

Besides the patch reconstruction method used, another consideration in example-based SR is the choice of the dictionary [18]. In [19], dictionary selection methods are divided into two classes: the analytic method and the learning-based method. The analysis dictionaries include Wavelets [20], Curvelets [21], Bandelets [22] and so on. These dictionaries have the advantage of fast implementation. The SR methods described above mainly adopt the learning-based dictionary, which is obtained from a set of training examples via the ML techniques and has a matrix representation. The matrix coefficients have good adaptability, which can be adapted to the training data. However, this dictionary is computationally expensive to apply. Rubinstein *et al.* [19] present a sparse dictionary learning model that is both adaptive and efficient.

Inspired by the idea of sparse dictionary model, we propose a fast example-based SR method in this paper. We reconstruct HR image patches using sparse dictionaries and the batch orthogonal matching pursuit (Batch-OMP) algorithm [23]. The Batch-OMP method is an efficient implementation of OMP algorithm. It can speed up the SR process, with compromising the reconstructed quality. Moreover, in the image reconstruction phase, an ANR model is introduced to further improve the algorithm efficiency. Experimental results and comparisons with other similar SR methods show that our algorithm can give competitive performance in terms of the image quality and running time.

The rest of this paper is organized as follows. Section 2 presents how to lean sparse dictionaries. Section 3 details the proposed SR method based on sparse dictionaries. Section 4 demonstrates the efficiency of this approach, and Section 5 concludes the paper.

2. Learning Sparse Dictionaries

The sparse coding model suggests that an input signal $x \in R^n$ can be represented as a linear combination of several specified atom signals, chosen from an over-complete dictionary $D \in R^{n \times N}$ (n < N). Given a set of training examples $X \in R^{n \times L}$, the problem of learning a dictionary for sparse coding can be formulated as [24]:

$$\min_{D,\Gamma} \left\| X - D\Gamma \right\|_F^2 \quad s.t. \quad \forall i, j \quad \left\| \gamma_i \right\|_0 \le t, \quad \left\| d_j \right\|_2 = 1, \tag{1}$$

where γ_i is the column of the sparse representation matrix Γ , d_j is the dictionary atom, and $\|\cdot\|_0$ denotes the number of non-zero entries of a vector. The model (1) can be solved by the K-SVD iteration algorithm [29].

Rubinstein *et al.* [19] propose a sparse dictionary model. And assume that each dictionary atom has itself a sparse representation over a specified base dictionary, *i.e.*

$$D = \Phi A, \tag{2}$$

where A is the atom representation matrix suggested to be sparse. Inspired by the original K-SVD algorithm, the sparse dictionary learning model is given by:

$$\min_{A,\Gamma} \left\| X - \Phi A \Gamma \right\|_{F}^{2} \quad \text{s.t.} \quad \begin{cases} \forall i \quad \left\| \gamma_{i} \right\|_{0} \leq t \\ \forall j \quad \left\| a_{j} \right\|_{0} \leq p, \quad \left\| \Phi a_{j} \right\|_{2} = 1 \end{cases} \tag{3}$$

with $\|a_j\|_0 \le p$ denotes the sparse constraint on D. Compared to analysis dictionaries and the model (1), the sparse dictionary model (3) can provide both adaptability via updating A, and fast implementation by choosing an appropriate base dictionary. Experimental results in [19] have shown the advantages of sparse dictionaries.

3. Proposed Methods

Image SR via sparse representation [12] aims to reconstruct HR images by jointly learning two coupled dictionaries, which enforces similar representation coefficients between the LR and HR image patch pair. In the learning phase, given a set of training image patch pairs $\{X_h, Y_l\}$, where Y_l is the corresponding LR patches of HR patches X_h . The dictionaries are obtained using the following minimization problem:

$$\{D_h, D_l, Z\} = \underset{D_h, D_l, Z}{\operatorname{arg\,min}} \frac{1}{N} \|X_h - D_h Z\|_F^2 + \frac{1}{M} \|Y_l - D_l Z\|_F^2 + \lambda (\frac{1}{M} + \frac{1}{N}) \|Z\|_1, \tag{4}$$

where D_h and D_l denote the HR and LR dictionaries respectively, N and M are the dimensions of HR and LR patches and Z is the sparse representation matrix. In the reconstruction phase, LR image patches are sparsely represented on the learned LR dictionary:

$$\alpha = \underset{\alpha}{\operatorname{arg\,min}} \|y_l - D_l \alpha\|_2^2 + \lambda \|\alpha\|_1, \qquad (5)$$

where y_l is an input LR feature patch, α is the sparse coefficient and λ is a balance parameter. Then the HR image patch can be estimated as $x = D_b \alpha$.

This paper builds upon the SR framework in [12] and proposes to reconstruct the HR image by learning sparse dictionaries. In addition, we introduce an ANR model on the sparse dictionaries to achieve fast SR.

3.1. Image SR Based on Sparse Dictionaries

Note that the joint dictionary learning model (4) does not guarantee that the sparse representation of a LR image patch can well reconstruct its underlying HR image patch. In order to make the learned dictionaries more compact, we train the HR and LR dictionaries in their own feature spaces. First, the sparse dictionary model (3) is applied to these LR image patches, resulting in the optimization of the LR dictionary:

$$\{A, \Gamma\} = \underset{A,\Gamma}{\operatorname{arg\,min}} \|Y_{l} - \Phi A \Gamma\|_{F}^{2} \quad \text{s.t.} \quad \begin{cases} \forall i \quad \|\gamma_{i}\|_{0} \leq t \\ \forall j \quad \|a_{j}\|_{0} \leq p, \quad \|\Phi a_{j}\|_{2} = 1 \end{cases}, \tag{6}$$

where the matrix A is the sparse representation of the dictionary, Γ is the coefficient matrix of the LR patches. The problem can be solved by alternating sparse coding and updating of A. As the HR image patches are recovered by the product of the HR dictionary and LR sparse coefficients. Then the HR dictionary D_h is

therefore constructed so that the approximation is as exact as possible. So the learning task for D_h is given by:

$$D_h = \underset{D_h}{\operatorname{arg\,min}} \left\| X_h - D_h \Gamma \right\|_F^2. \tag{7}$$

The solution of this problem is given by the following Pseudo-Inverse expression:

$$D_h = X_h \Gamma^{\dagger} = X_h \Gamma^T \left(\Gamma \Gamma^T \right)^{-1}. \tag{8}$$

Thus far, we have obtained the sparse representation matrix A of the LR dictionary and the HR dictionary D_h . In the reconstruction phase, for the input LR patch y_l , its sparse representation is calculated by the following optimization:

$$\gamma = \underset{\gamma}{\operatorname{arg\,min}} \| y_l - \Phi A \gamma \|_2^2 \quad \text{s.t.} \quad \| \gamma \|_0 \le t, \tag{9}$$

Then the corresponding HR patch is derived as $x = D_b \gamma$.

Moreover, we use the Batch-OMP algorithm for sparse coding, which can achieve good improvements for fast implementation compared to the traditional OMP method. Our experimental results show that the dictionary learning model (6) and reconstruction model (9) can be solved efficiently while maintaining the reconstructed quality.

3.2. Anchored Neighborhood Regression

Timofte et al. [17] use the ridge regression and reformulate the problem (5) as:

$$\gamma = \arg\min_{\alpha} \left\| y_l - D_l \gamma \right\|_2^2 + \lambda \left\| \gamma \right\|_2. \tag{10}$$

The closed-form solution is given by:

$$\gamma = \left(D_l^T D_l + \lambda I\right)^{-1} D_l^T y_l.$$

The HR patch can then be calculated using the coefficient and D_h

$$x = D_h \gamma = D_h \left(D_l^T D_l + \lambda I \right)^{-1} D_l^T y_l, \tag{11}$$

where the projection matrix $P=D_h\left(D_l^TD_l+\lambda I\right)^{-1}D_l^T$ can be computed offline. Hence, it only needs multiply the input LR feature patch and P to construct the HR patch.

The above SR approach can be regarded as a regression on the entire dictionary. Instead, the ANR algorithm [17] aims to calculate a separate projection matrix P_j for each dictionary atom d_j by employing its K nearest neighbors as the dictionaries $\left\{D_{_h}^j, D_{_l}^j\right\}$. The SR problem can then be solved by finding the nearest dictionary atom d_j for each input LR patch y_l , followed by calculating the HR patch using the corresponding projection matrix P_j :

$$x = P_i y_i$$

[17] adopts a learned dictionary as the approach of Zeyde *et al.* [13]. Here, we combine the sparse dictionaries $\{D_h, D_l\}$ in Section 3.1 with the ANR method to achieve the fast SR. In the next section, the experimental results will show the comparisons of different methods in terms of the running time and recovery performance.

4. Experimental Results

We refer to the proposed method based on the sparse dictionaries in Section 3.1 as our method (i), and the method combined with the ANR model in Section 3.2 as our method (ii). The proposed methods are compared with Bicubic interpolation, the representative sparse coding-based methods of Yang *et al.* [12] and Zeyde *et al.* [13], and ANR [17]. In our experiments, we magnify the LR image by a factor of 3. We use the same training data as Zeyde *et al.* [13]. The HR feature patches are subtracted means, and the LR patches are derived by concatenating the first and second order gradients. The Structural Similarity Index (SSIM) [25] and PSNR are employed to evaluate the objective quality of the proposed and comparison methods. All the experiments were tested using MATLAB on an Intel E8400, 3.0 GHz, 3 G RAM.

In the sparse dictionary learning phase, the image patch size is 5×5 and around 130000 training patch-pairs are collected. We apply the over-complete DCT as the based dictionary, fix 1000 atoms in the dictionary, and use p=16 coefficients per atom. We set the initial matrix A to identity, the maximum number of iterations as 40, and t=3 atoms per patch representation. In our experiments, the average time of the proposed method for learning dictionaries is 8 minutes, which of Yang $et\ al.$ [12] and Zeyde $et\ al.$ [13] are 4 hours and 12 minutes, respectively. This mainly owes to the application of the Batch-OMP algorithm.

Table 1. PSNR (dB) and Running Time (s) Results of Different Methods for The Test Images (Magnification ×3)

Images	Bicubic		Yang et al. [12]		Zeyde et al.[13]		Our method (i)		ANR [17]		Our method (ii)	
	PSNR	Time	PSNR	Time	PSNR	Time	PSNR	Time	PSNR	Time	PSNR	Time
baboon	23.2	0.0	23.5	136.1	23.5	5.7	23.6	3.9	23.6	1.7	23.6	1.7
barbara	26.2	0.0	26.4	163.8	26.8	10.2	26.8	7.1	26.7	3.1	26.8	3.2
bridge	24.4	0.0	24.8	160.3	25.0	6.2	25.2	4.4	25.0	1.9	25.1	1.9
coastguard	26.6	0.0	27.0	39.7	27.1	2.3	27.1	1.6	27.1	0.7	27.1	0.7
comic	23.1	0.0	23.9	54.0	24.0	2.1	24.1	1.5	24.0	0.7	24.0	0.6
face	32.8	0.0	33.1	21.9	33.5	1.7	33.5	1.2	33.6	0.5	33.5	0.5
flowers	27.2	0.0	28.2	80.0	28.4	4.2	28.4	3.0	28.5	1.3	28.3	1.2
foreman	31.2	0.0	32.0	27.3	33.2	2.4	33.3	1.6	33.2	0.7	33.1	0.8
lenna	31.7	0.0	32.6	78.4	33.0	6.2	33.0	4.4	33.1	1.9	32.9	2.0
man	27.0	0.0	27.8	111.9	27.9	6.1	28.0	4.2	27.9	1.8	28.0	1.8
monarch	29.4	0.0	30.7	118.5	31.1	9.3	30.9	6.4	31.1	2.7	30.9	2.8
pepper	32.4	0.0	33.3	67.9	34.1	6.2	33.9	4.3	33.8	1.9	33.8	1.7
ppt3	23.7	0.0	25.0	100.0	25.2	7.5	25.1	5.1	25.0	2.4	25.0	2.3
zebra	26.6	0.0	28.0	115.2	28.5	5.4	28.6	3.8	28.4	1.6	28.5	1.6
average	27.54	0.02	28.31	91.09	28.67	5.39	28.68	3.75	28.65	1.63	28.61	1.63

In the reconstruction phase, we chose 14 test images, and the results of Bicubic interpolation as the initial HR images. Table 1 shows the PSNR values and running time of different methods for the test images. It shows that our methods outperform the method of Yang *et al.*, and are competitive with the current state-of-the-art ANR algorithm. Although the proposed method (i) performs slightly worse than the method of Zeyde *et al.* for PSNR, it speeds up the SR process by 30%. The proposed method (ii) gets similar quality performance as the method (i), but its running time reduces by 43% due to the introduce of the ANR model.



Figure 1. Reconstructed HR Images and SSIM Values of Barbara by Different Methods (a) The LR Image; (b) The Original Image; (c) Bicubic (0.651); (d) Yang [12] (0.694); (e) Zeyde [13] (0.698); (f) Our Method (i) (0.705); (g) ANR [17] (0.703); (h) Our Method (ii) (0.704)

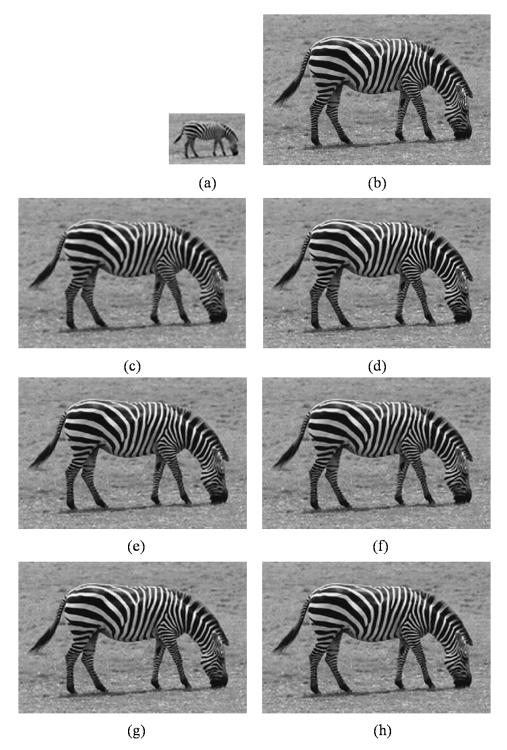


Figure 2. Reconstructed HR Images and SSIM Values of Zebra by Different Methods (a) The LR Image; (b) The Original Image; (c) Bicubic (0.795); (d) Yang [12] (0.827); (e) Zeyde [13] (0.842); (f) Our Method (i) (0.843); (g) ANR [17] (0.843); (h) Our Method (ii) (0.841)

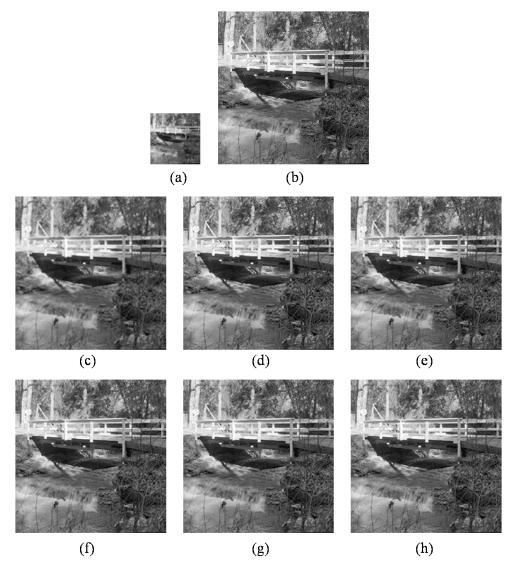


Figure 3. Reconstructed HR Images and SSIM values of Bridge by Different Methods (a) The LR Image; (b) The Original Image; (c) Bicubic (0.611); (d) Yang [12] (0.635); (e) Zeyde [13] (0.650); (f) Our Method (i) (0.653); (g) ANR [17] (0.654); (h) Our Method (ii) (0.654)

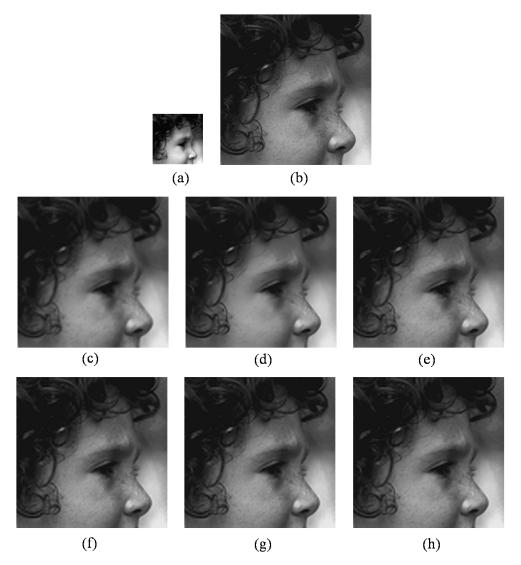


Figure 4. Reconstructed HR Images and SSIM values of Face by Different Methods (a) The LR Image; (b) The Original Image; (c) Bicubic (0.797); (d) Yang [12] (0.800); (e) Zeyde [13] (0.819); (f) Our Method (i) (0.821); (g) ANR [17] (0.822); (h) Our Method (ii) (0.821)

In Figures 1-4, we show some reconstructed HR images by the above algorithms. From Figures 1-3, we can see that HR images by Bicubic interpolation have blurred edges and artifacts. The method of Yang *et al.* shows improvements in most places, but fails in the areas with fine image details (*e.g.*, the cloth in Figure 1(d) and the face in Figure 4(d)). By contrast, our methods show better performance both in terms of objective metrics and subjective visual results. This is mainly because that the sparse dictionaries learning approach can improve the reconstruction accuracy. Compared to the method of Zeyde *et al.* and ANR, our methods can achieve competitive visual quality and SSIM values.

From Table 1 and Figures 1-4, we can see that the proposed methods get visually pleasant results with high objective index values. Moreover, it can achieve fast SR. This clearly demonstrates the effectiveness of sparse dictionaries in solving the problem of fast image SR.

5. Conclusion

In this paper, we propose a fast example-based SR method, which introduced a novel dictionary learning approach. The approach employs a base dictionary to learn the sparse matrix of the dictionary, being adaptive and efficient. The Batch-OMP algorithm is used to sparse coding and can provide great improvement in calculation speed. Moreover, we incorporate the ANR model into the sparse dictionaries and achieve fast SR using the projection matrices. Simulation results indicate that the proposed method can obtain qualitatively reconstructed images with high values of PSNR and SSIM. In future work, we would like to learn an internal dictionary from the given image online to obtain better adaptivity and reconstruction.

Acknowledgments

This work is supported by creation fund of Jinan University and the National Natural Science Foundation of China under Grant no. 61401368. The authors would like to thank authors of Refs. [12, 13, 17] for providing their programs online.

References

- [1] X. Li and M. T. Orchard, "New edge-directed interpolation", IEEE Transactions on Image Processing, vol. 10, no. 10, (2001), pp. 1521-1527.
- [2] Q. Zhang, J. Wu, "Image super-resolution using windowed ordinary Kriging interpolation", Optics Communications, vol. 336, (2015), pp. 140–145.
- [3] H. Kim, Y. Cha, S. Kim, "Curvature interpolation method for image zooming", IEEE Transactions on Image Processing,vol.20, no. 7, (2011), pp.1895--1903.
- [4] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super-resolution", IEEE Transactions on Image Processing, vol. 13, no. 10, (2004), pp.1327–1344.
- [5] H. Zhang, Y. Zhang, H. Li, and T. Huang, "Generative Bayesian Image Super Resolution With Natural Image Prior", IEEE Transactions on Image Processing, vol. 21, no. 9, (2012), pp.4054-4067.
- [6] W. T. Freeman, E. C. Pasztor and O. T. Carmichael. "Learning low-level vision", International Journal of Computer Vision, vol. 40, no. 1, (2000), pp. 25-47.
- [7] H. Chang, D.-Y. Yeung and Y. Xiong, "Super-resolution through neighbor embedding", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, (2004).
- [8] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding", Science, vol. 290, no. 5500, (2000), pp. 2323-2326.
- [9] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. Alberi Morel, "Low-Complexity Single-Image Super-Resolution based on Nonnegative Neighbor Embedding", In BMVC, (2012).
- [10] X. Gao, K. Zhang, D. Tao and X. Li, "Image super-resolution with sparse neighbor embedding", IEEE Transaction on Image Processing, vol. 21, no. 7, (2012), pp. 3194-3205.
- [11] Y. Wang, J. Fan, J. Xu, and X. Wu, "Learning-based Super-resolution via Canonical Correlation Analysis", International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 8, no. 6, (2015), pp.69-84.
- [12] J. Yang, J. Wright, T. Huang and Y. Ma, "Image super-resolution via sparse representation", IEEE Transaction on Image Processing, vol. 19, no. 11, (2010), pp. 2861-2873.
- [13] R. Zeyde, M. Elad and M. Protter, "On single image scale-up using sparse-representation", Springer Berlin Heidelberg 7th International Conference Curves and Surfaces, Avignon, France, (2010), June 24-30.
- [14] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization", IEEE Transaction on Image Processing, vol. 20, no. 7, (2011), pp. 1838–1857.
- [15] K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 32, no. 6, (2010), pp. 1127-1133.
- [16] J. Yang, Z. Lin and S. Cohen, "Fast image super-resolution based on in-place example regression", IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, (2013), June 23-28.
- [17] R. Timofte, V. D. Smet and L. V. Fool, "Anchored neighborhood regression for fast example-based super-resolution", IEEE International Conference on Computer Vision, Sydney, NSW, (2013), December 1-8.

- [18] M. Bevilacqua, A. Roumy and M.-L Morel. "A Single-Image Super-Resolution via Linear Mapping of Interpolated Self-Examples", IEEE Transaction on Image Processing, vol. 23 no. 12, (2014), pp. 5334-5347.
- [19] R. Rubinstein, M. Zibulevsky and M. Elad, "Double Sparsity: Learning Sparse Dictionaries for Sparse Signal Approximation", IEEE Transaction on Image Processing, vol. 53, no. 3, (2010), pp. 1553-1564.
- [20] S. Mallat, "A wavelet tour of signal processing", Academic press, (1999).
- [21] E. J. Candès and D. L. Donoho, "Curvelets— A Surprisingly Effective Nonadaptive Representation for Objects With Edges", Nashville, TN: Vanderbilt University Press, (1999), pp. 105-120.
- [22] E. LePennec and S. Mallat, "Sparse geometric image representations with bandelets", IEEE Transaction on Image Processing, vol. 14, no. 4, (2005), pp. 423-438.
- [23] R. Rubinstein, M. Zibulevsky and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit", Technical Report CS Technion, (2008).
- [24] M. Aharon, M. Elad, and A. M. Bruckstein, "The K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation", IEEE Transaction on Signal Processing, vol. 54, no. 11, (2006), pp. 4311–4322.
- [25] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity", IEEE Transaction on Image Processing, vol. 13, no. 4, (2004), pp. 600–612.

International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 10, No. 6 (2017)