

Image super-resolution via sparse representation over multiple learned dictionaries based on edge sharpness and gradient phase angle

F. Yeganli¹ · M. Nazzal¹ · H. Ozkaramanli¹

Received: 3 January 2015 / Revised: 11 June 2015 / Accepted: 7 September 2015 / Published online: 12 October 2015
© Springer-Verlag London 2015

Abstract This paper introduces an algorithm for single-image super-resolution based on selective sparse representation over a set of low- and high-resolution cluster dictionary pairs. Patch clustering in the dictionary training stage and model selection in the reconstruction stage are based on patch sharpness and orientation defined via the magnitude and phase of the gradient operator. For each cluster, a pair of coupled low- and high-resolution dictionaries is learned. In the reconstruction stage, the most appropriate dictionary pair is selected for the low-resolution patch and the sparse coding coefficients with respect to the low-resolution dictionary are calculated. A high-resolution patch estimate is obtained by multiplying the sparse coding coefficients with the corresponding high-resolution dictionary. The performance of the proposed algorithm is tested over a set of natural images. Results validated in terms of PSNR, SSIM and visual comparison indicate that the proposed algorithm is competitive with the state-of-the-art super-resolution algorithms.

Keywords Single-image super-resolution · Sparse representation · Dictionary learning · Sharpness measure · Coupled dictionaries · Gradient phase angle

1 Introduction

Single-image super-resolution (SR) is an ill-posed inverse problem of reconstructing a high-resolution (HR) image estimate based on a low-resolution (LR) image of the same scene. Several approaches to SR exist in the literature. These can be classified into three main categories [1,2]. The first category includes interpolation-based methods where unknown pixel values are estimated by interpolation. Interpolation methods tend to blur the SR outcome. However, interpolation is shown to perform better when combined with exploiting natural image features. Still, interpolation methods lack the capability of handling the visual complexity of natural images, especially for fine textures. The second category is reconstruction-based methods which impose reconstruction constraints on the HR image estimation. Such constraints enforce a similarity between a blurred and downsampled version of the HR image estimate and the LR image. However, such methods produce jaggy or ringing artifacts around edges because of the lack of regularization. The third category is the learning-based methods employing a training stage and a testing stage. This is based on utilizing the relationship between the LR and HR image patches as a natural image prior. This is carried out by assuming a similarity between training and testing sets of signals. One of the most successful learning approaches is the sparse representation-based approach [1,2].

In the context of sparse representation SR, sparsity is effectively used as a regularizer. This usage is essentially based on the assumed invariance of the sparse coding coefficients of HR image patches and their LR counterparts with respect to scale, as originally proposed by Yang et al. in [1] and [3]. This means the possibility of reconstructing a HR patch by multiplying the HR dictionary with the sparse coding coefficients of its LR counterpart.

✉ F. Yeganli
yeganli@gmail.com; faezeh.yeganli@emu.edu.tr
M. Nazzal
mahmoud.nazzal@emu.edu.tr
H. Ozkaramanli
huseyin.ozkaramanli@emu.edu.tr

¹ Electrical and Electronic Engineering Department, Eastern Mediterranean University, Via Mersin 10, Gazimagusa, TRNC, Turkey

The sparse coding problem requires the availability of a dictionary. It is customary to obtain dictionaries by training over example signals through the dictionary learning (DL) process. In the DL process, a dictionary is learned aiming at giving a sparse yet loyal representation to signals in a given training set. The fundamental advantage of learned dictionaries is the signal fitting capability [4]. It is well known that the representation power of a learned dictionary depends on its redundancy. However, high redundancy of a dictionary tends to cause instabilities and degrade the representation quality [5,6].

A recent research trend considers the design of class-dependent dictionaries. The basic motivation behind this approach is that signal variability within a class is much less than the general signal variability. Along this line, Dong et al. [7] use K-means clustering to divide the training data into a number of clusters, and learn compact cluster dictionaries. They then adaptively select the most relevant cluster dictionary for a given signal. Feng et al. [8] split the signal space into subspaces using K-space clustering and extract shared bases in these subspaces to form a dictionary. More recently, attention has been paid toward designing structural dictionaries. An attempt at this is the work of Yu et al. in [9] where they used a composition of orthogonal basis functions to construct a dictionary. Each basis is concerned with a specific signal structure. They applied a corresponding structural sparse model selection over the structural dictionary. In [10], Yang et al. proposed a multiple geometric dictionary-based clustered sparse coding scheme, which first trains the geometric dictionaries of geometric clusters, and then image patches are sparsely coded over different dictionaries.

Sharp image regions contain high-frequency features. Reconstruction of such regions is a challenging part for SR algorithms. Therefore, representation of these features is crucial for the SR problem. It is noted that image regions having high-frequency details have high magnitudes of the gradient operator. In this regard, one can define a sharpness measure (SM) [11] as an indicator to image spatial intensity variations. The gradient operator has been employed as a natural image prior to solve ill-posed image processing problems such as SR [12] and denoising [13].

In [14], we proposed a SR algorithm based on sparse representation over multiple learned dictionaries. SM is shown to possess approximate scale invariance and is thus used as a clustering and model selection criterion. In [14], SM is defined in terms of the magnitude of the patch gradient operator and is used to classify patches based on their spatial intensity variations. In this paper, we extend the work conducted in [14] making use of the patch gradient operator's phase as a directional secondary classifier. The scale invariance of this classifier has been established in [15,16]. Such a classifier is defined in terms of the dominant phase angle (DPA) of the gradient operator. In this work, three main SM

clusters are employed. Then, data in each SM cluster are further clustered into five DPA sub-clusters. In this setting, sub-clusters have distinct sharpness levels and directionality. For each sub-cluster, a dictionary pair is learned in the DL stage. In the reconstruction stage, the SM and DPA values of each LR patch are used to classify it into a certain sub-cluster. Then, the sparse coding coefficients of this patch with respect to the cluster LR dictionary are calculated. Then, a HR patch is reconstructed by multiplying the cluster HR dictionary with the calculated coefficients.

Experiments conducted on natural images validate a competitive performance of the proposed algorithm as compared to the state-of-the-art SR algorithms at reasonable computational complexity. This result is validated in terms of the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) quality measures as well as visual comparisons.

The remainder of this paper is organized as follows. Section 2 details the single-image SR via sparse representation approach. In Sect. 3, the proposed SR algorithm is presented. Section 4 presents experimental results testing the performance of the proposed algorithm and the representation power of the learned dictionaries. In Sect. 5, conclusions are made.

2 Single-image super-resolution via sparse representation

Denoting by x_H a HR image patch reshaped into the column form, one can write a sparse approximation of this patch over a HR dictionary D_H with a sparse representation coefficient vector α_H , as follows

$$x_H \approx D_H \alpha_H, \quad (1)$$

In the same manner, given a dictionary trained over LR image patches as D_L , a LR patch of the same scene x_L can be sparsely represented with a sparse representation coefficient vector α_L , as follows

$$x_L \approx D_L \alpha_L. \quad (2)$$

It is customary to model the relationship between x_H and x_L via a blurring and downsampling operator Ψ , as $x_L \approx \Psi x_H$. Given that D_L and D_H are trained in a coupled manner, one can further assume that the same operator links D_L and D_H , as $D_L \approx \Psi D_H$. Following this assumption, it can be shown that $x_L \approx \Psi x_H \approx \Psi D_H \alpha_H \approx D_L \alpha_H$, or equivalently, $\alpha_H \approx \alpha_L$. This forms the foundation for HR patch reconstruction. More precisely, given D_H and the sparse coding vector of x_L over D_L as α_L , one may reconstruct the corresponding HR patch as

$$x_H \approx D_H \alpha_L. \quad (3)$$

3 The proposed super-resolution algorithm

In this work, SM of image patches is first used to classify image patches into three main clusters based on their sharpness. Then, DPA is used to further cluster the patches of each main cluster into several sub-clusters based on their directional structure. Analogously, these two measures are employed to select the most relevant cluster dictionary pair for each LR patch during the reconstruction stage.

3.1 Clustering and sparse model selection with the patch sharpness measure and dominant phase angle

The image gradient operator is believed to be crucial to the perception and analysis of natural images [17–19]. SM can be defined in terms of the magnitude of the gradient operator as a numerical measure to quantify the spatial intensity variations of image patches. For a given image patch, SM is defined as [20]

$$SM = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \sqrt{|G_{i,j}^h| + |G_{i,j}^v|}, \quad (4)$$

where G^h and G^v denote the horizontal and vertical gradients and N_1 and N_2 denote the patch dimensions.

In [12, 21], Sun et al. defined the gradient profile prior as a 1D profile of gradient magnitudes perpendicular to image structures, and studied its behavior with respect to scale. They reported that the edge sharpness of natural images has a statistical distribution which is independent of scale. The invariance of SM is shown to be stronger for image patches that contain strong edges, corners and texture. In this work, we employ three SM clusters C_1 , C_2 and C_3 defined, respectively, for SM intervals of [0, 10], [10, 20] and [20, 255]. The lower bound of the third interval is set in such a way that it contains very sharp patches. The bounds of the other two cluster intervals are set to uniformly divide the remaining SM range. In this setting, patches are classified as unsharp, moderately sharp and very sharp. Finer classification can be obtained by defining more clusters. However, this will deteriorate the ability of SM in correctly estimating which cluster a given patch belongs to.

It is well acknowledged that the phase of a quantity is generally more informative than its magnitude. Accordingly, it seems advantageous to think about exploiting information in the phase of the gradient operator [20]. It has been shown that the histogram of gradient phase angles has an acceptable degree of scale invariance in [10]. One can define the phase matrix of the gradient operator based on the horizontal and vertical gradients [20] as follows

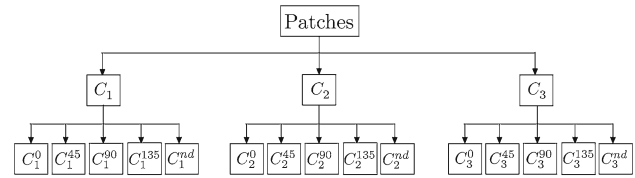


Fig. 1 The proposed 2-level clustering scheme

$$\Phi_{i,j} = \arctan\left(\frac{G_{i,j}^v}{G_{i,j}^h}\right). \quad (5)$$

The phase matrix Φ can be used to determine the orientation of a given patch. This can be done by inspecting the histogram of the angles in Φ after quantizing them to a set of values. If the histogram significantly peaks at a certain value, one may assume that this value describes the dominant directional nature of the patch. However, if the histogram is flat, this means that the patch is either smooth or is composed of several directions. In both cases, it can be assumed to be non-directional. In this work, we quantize the angles in Φ into values of 0° , 45° , 90° and 135° . Then, if a certain angle in Φ is repeated more than 50%, we assume that this angle is a dominant angle. Otherwise, we assume that the patch is non-directional. Figure 1 shows the two-stage clustering based on SM and DPA. In total, 15 sub-clusters are defined. These are denoted by C_1^0 through C_3^{nd} , where the subscript denotes the SM cluster and the superscript represents the DPA cluster (*nd* stands for non-directional).

The proposed algorithm is composed of two stages. The first one is the training stage, where a set of dictionary pairs is trained. The second one is the reconstruction stage where the best dictionary pair is selected to sparsely reconstruct HR patches from their LR counterparts.

The training stage requires a set of HR images along with the corresponding LR images. A LR image is obtained by blurring and downsampling the HR one. Each LR image is then interpolated by a scale factor of 2 to the dimensions of the corresponding HR image. This is referred to as the middle-resolution (MR) image. Then, feature extraction filters are applied to the MR images to extract features as done in [1]. Dictionary learning and sparse coding of the LR patches is done with these features. This is shown to be more advantageous than dealing with LR patches directly [1, 3].

LR and HR patches corresponding to the same spatial location are handled as pairs. Each MR patch is then classified into a specific cluster based on its SM and DPA values. The HR patch in the pair is placed into the same cluster. The mean value of each HR patch is subtracted to allow for better dictionary learning. LR and HR patches of each cluster are used to train for a pair of coupled LR and HR cluster dictionaries, respectively. For this purpose, the method proposed in [1] is used. Algorithm 1 outlines the main steps of the training stage.

Algorithm 1 The Proposed Cluster DL Algorithm.

- 1: **INPUT:** HR Training Image Set.
- 2: **OUTPUT:** A Set of Directional Cluster Dictionary Pairs with Varying Sharpness.
- 3: Divide each HR image into patches and subtract the mean value of each patch.
- 4: Reshape patches into vectors and combine them column-wise to form a HR training array.
- 5: Blur (bicubic kernel) and downsample each HR image to generate a LR image.
- 6: Divide each LR image into patches.
- 7: Upsample each LR image to the MR level.
- 8: Apply feature extraction filters on each MR image.
- 9: Divide the extracted MR features into patches and reshape them into column vectors.
- 10: Combine the MR features column-wise to form the LR training array.
- 11: **for** Each patch in the LR training array, **do**
- 12: Calculate SM and DPA of the MR patch, and identify the corresponding cluster.
- 13: Add the MR patch to the cluster LR training set.
- 14: Add the corresponding HR patch to the HR training set of this cluster.
- 15: **end for**
- 16: For each cluster, learn a pair of coupled dictionaries.

Algorithm 2 The Proposed Single-Image SR Algorithm.

- 1: **INPUT:** A LR Test Image, Cluster Dictionary Pairs.
- 2: **OUTPUT:** A HR Image Estimate
- 3: Divide the LR image into overlapping patches.
- 4: Upsample the LR image to MR level.
- 5: Extract features from the MR image.
- 6: Divide the extracted MR features into overlapping patches and reshape them into vectors.
- 7: **for** Each MR patch **do**
- 8: Calculate SM and DPA of the MR patch.
- 9: Determine the cluster this patch belongs to.
- 10: Sparsely code the features of the MR patch over the cluster LR dictionary.
- 11: Reconstruct the corresponding HR patch by right-multiplying the HR dictionary of the same cluster with the sparse codes of the MR features.
- 12: **end for**
- 13: Reshape and merge overlapping patches to obtain a HR image estimate.

In the reconstruction stage, a LR image is first upsampled using bicubic interpolation to the MR level. Features are extracted by applying feature extraction filters [1] and then reshaped into the vector form. A certain patch overlap is allowed to assure local consistency between the reconstructed patches [1]. The SM and DPA values of each MR patch are calculated and the cluster that the MR patch belongs to is identified. Using the dictionary pair of the identified cluster, first the sparse representation coefficient vector of the corresponding MR feature vector over the cluster LR dictionary is calculated. Then the HR patch is reconstructed by right-multiplying the cluster HR dictionary with the calculated sparse representation coefficient vector. Finally, a HR image is obtained by reshaping the reconstructed HR patches

into the two-dimensional form and merging them. The proposed reconstruction algorithm is detailed in Algorithm 2.

As compared to the standard SR algorithm of Yang et al. [1] in terms of computational complexity, the proposed algorithm has two overheads. First, it uses multiple dictionary pairs instead of one. Second, it calculates the SM and DPA values of every patch for the purpose of clustering and model selection. However, the proposed algorithm can be implemented while significantly reducing the computational complexity without degrading the performance quality. The fact that a cluster corresponds to a specific signal class makes it possible to learn more compact cluster dictionaries. Since the computational complexity of sparse coding depends on the dictionary dimensions, using more compact dictionaries will substantially reduce this complexity. Moreover, it is noted that many patches are located in C_1 . This cluster contains patches with insignificant high-frequency components. Accordingly, one may afford to use bicubic interpolation to reconstruct patches in C_1 without degrading the reconstruction quality. This means that a large percentage of patches (located in C_1) require only calculating SM values and applying bicubic interpolation. Therefore, the computational complexity of the proposed algorithm is comparable to that of Yang et al.'s algorithm in [1].

4 Experimental validation

In this section, the performance of the proposed algorithm is examined and compared to several SR algorithms. These include the basic sparse representation-based SR algorithm of Yang et al. [1] which employs one dictionary pair as well as the SR algorithms of Peleg et al. [22] and He et al. [23] as state-of-the-art techniques. These algorithms are different in nature. For comparisons, care has been taken to ensure that the parameters used in the training and testing stages of these algorithms are as close to each other as possible. If a parameter is unique to a specific algorithm, the value suggested by the authors is used. Image SR results for a scale factor of 2 are presented. However, the proposed algorithm can easily be modified for other scale factors.

Test images include some well-known benchmark images which were used in [22–24]. Several other images have also been selected from different datasets [25–27] because of their rich high-frequency contents. All test images are shown in Fig. 2.

Dictionaries of the proposed algorithm are learned as specified in Algorithm 1. Dictionary training for the proposed algorithm is done over the 1000-image Flickr dataset [28], along with some typical text images. These text images are added to the training set to be sure about the availability of enough patches with relatively high SM values. The clustering of LR and HR training patches is carried out in terms of



Fig. 2 Test images from left to right and top to bottom: Barbara, BSDS 198054, building image 1, build poster, butterfly, fence, house, Kodak 08, Lena, ppt3, text image 3 and texture

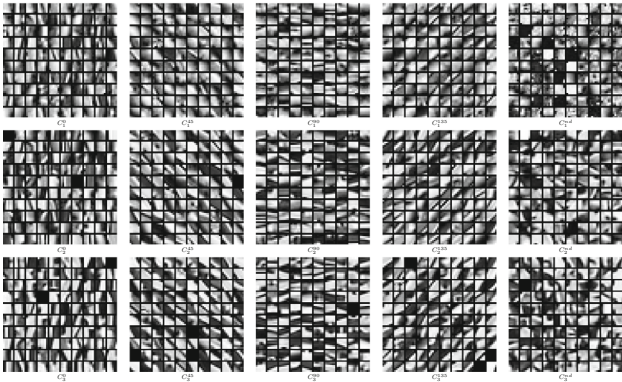


Fig. 3 Reshaped example atoms of the 15 HR cluster dictionaries

the SM and DPA of each LR patch, as described in Sect. 3.1. We then randomly selected 40,000 pairs of LR and HR training patches for each cluster; 600-atom dictionary pairs are designed for the proposed algorithm.

Example reshaped atoms of HR cluster dictionaries designed by the proposed algorithm are shown in Fig. 3. It can be clearly observed that the designed dictionaries inherit the sharpness nature and directionality of their respective clusters. It is notable that the dictionaries in C_1 are not sharp, whereas these in C_2 are moderately sharp and the dictionaries in C_3 are sharper.

For the algorithm of Yang et al. [1], a single dictionary pair with 1000 atoms is learned. For the learning, 40,000 LR and HR patch pairs are randomly selected from the same training set used by the proposed algorithm. We used the default design parameters and training image datasets for the algorithms of Peleg et al. [22] and He et al. [23] as specified by the authors.

It is noted that in case of color images, SR is applied only to the luminance color channel, whereas the chrominance color channels are reconstructed with bicubic interpolation, as customarily done in most SR algorithms. The three components are used to obtain a full-color HR image. In these experiments, we employ PSNR [29] as a quantitative measure of quality. For color images, PSNR is calculated with the luminance color components of the original image and the reconstructed image, in accordance with the common practice in the literature.

Also, SSIM [29] is used as a perceptual quality metric, which is believed to be more compatible with human perception than PSNR. Similar to most SR algorithms, we calculate SSIM for color images as the average SSIM value of the luminance and two chrominance components of the image.

Before presenting the simulation results, we establish the discrimination power of the designed dictionaries. In other words, we establish that data in a given cluster is best represented with the dictionary pair designed for that cluster. For this purpose, we use the training patch pairs of each cluster as testing signals. Then we reconstruct HR patches in each cluster using all of the fifteen dictionary pairs. For comparison, the HR patch in each cluster is also reconstructed with the single dictionary pair of Yang et. al's algorithm [1] and bicubic interpolation. The mean squared error (MSE) between the ground-truth and reconstructed HR patches is recorded, and results are presented in Table 1. It is clearly observed that data in each cluster are best reconstructed with the dictionary pair designed specifically for that cluster. One exception is the cluster C_1^{nd} where bicubic interpolation method produces slightly lower MSE. It is observed that the error level in the first SM cluster C_1 is not significantly different for all methods. This is because data in this cluster are unsharp. Thus, patches in sub-clusters of C_1 (C_1^0 through C_1^{nd}) can be satisfactorily reconstructed with bicubic interpolation. Due to the fact that the majority of patch pairs belong to cluster C_1 , the employment of bicubic interpolation also serves to reduce the computational complexity of the SR without degrading the reconstruction quality. However, error levels for sharper clusters C_2 and C_3 are significantly higher. The difference in MSE of different methods is also significantly higher.

Since the proposed algorithm uses multiple dictionary pairs, there is an inherent trade-off between model selection and reconstruction quality. Model selection relies on the SM and DPA values of the LR patch which are approximately scale invariance. It is then assumed that the HR patch has similar SM and DPA values in view of this invariance. The validity of this assumption still needs more investigation.

We now carry out another experiment to determine the average reconstruction quality when the correct model selection is made with the designed dictionaries. Four scenarios are described, and the average PSNR and SSIM results are recorded in Table 2. In the first scenario (S_1), SM is used alone for both clustering in the training stage and model selection in the reconstruction stage. Three main clusters with SM intervals of $[0, 10]$, $[10, 20]$ and $[20, 255]$ are considered. A 1000-atom dictionary pair is learned for each cluster. The second scenario (S_2) is exactly the same as S_1 , but with employing perfect model selection. In this context, perfect model selection is carried out by super-resolving each LR patch with each of the three cluster dictionary pairs. Then the super-resolved patch that is closest to the ground-truth HR patch in the MSE sense is selected. In the third scenario

Table 1 Reconstruction MSE for the patches in the 15 clusters with the dictionary pair of each cluster, the single dictionary pair of Yang et al. (SD) and with bicubic interpolation

In cluster	MSE with																
	Dictionary pairs of C_1					Dictionary pairs of C_2					Dictionary pairs of C_3					SD	Bicubic
	C_1^0	C_1^{45}	C_1^{90}	C_1^{135}	C_1^{nd}	C_2^0	C_2^{45}	C_2^{90}	C_2^{135}	C_2^{nd}	C_3^0	C_3^{45}	C_3^{90}	C_3^{135}	C_3^{nd}		
C_1^0	70.87	83.70	87.45	82.68	84.38	73.77	84.96	86.88	83.14	76.76	77.42	84.88	90.98	85.28	77.86	80.16	75.50
C_1^{45}	42.67	33.95	42.66	46.08	42.32	43.71	35.54	42.80	48.45	38.87	46.84	38.88	47.23	51.19	41.28	40.89	37.09
C_1^{90}	53.12	49.04	42.16	48.68	48.97	54.35	50.11	43.83	49.49	45.16	55.94	52.26	46.24	51.66	46.74	48.81	43.67
C_1^{135}	46.23	50.25	45.54	36.53	46.52	46.29	51.98	46.11	38.39	41.41	50.73	55.02	50.67	41.22	44.01	43.66	40.44
C_1^{nd}	18.45	18.25	18.07	18.56	16.19	18.60	18.31	18.10	18.68	17.05	19.77	19.68	20.08	19.59	18.78	18.50	14.10
C_2^0	294.91	357.10	360.18	348.95	348.24	273.24	349.12	352.24	340.46	301.14	284.21	341.37	363.30	342.47	295.02	319.54	322.75
C_2^{45}	244.56	189.00	241.04	267.02	241.97	238.61	174.87	232.78	273.36	197.22	247.26	186.62	249.62	281.14	202.75	214.17	226.96
C_2^{90}	401.01	375.84	330.03	375.56	370.40	402.52	379.46	300.58	374.25	327.49	401.48	382.15	315.81	380.23	330.78	356.02	356.68
C_2^{135}	244.08	270.39	244.61	189.59	242.92	234.20	271.63	235.03	175.64	197.54	249.78	283.57	248.34	186.83	204.56	215.05	227.32
C_2^{nd}	455.99	459.41	454.35	457.50	455.56	445.35	456.67	440.76	449.47	389.45	456.52	462.21	462.67	457.80	413.17	442.64	455.15
C_3^0	772.92	1056.32	983.82	995.49	956.85	704.33	1006.21	986.82	943.57	797.27	567.06	954.01	1000.68	937.18	675.07	808.46	930.51
C_3^{45}	715.85	484.46	686.26	783.87	699.11	670.10	426.21	653.14	812.45	489.63	677.12	372.02	670.83	816.85	464.28	535.95	664.77
C_3^{90}	1051.33	993.12	798.80	988.96	941.99	1053.50	999.48	727.89	975.10	783.75	1041.49	983.84	609.26	974.84	750.02	853.35	947.18
C_3^{135}	698.56	782.26	702.75	478.65	691.11	650.64	789.91	649.97	423.39	502.11	674.45	829.74	673.11	371.36	470.26	539.85	663.32
C_3^{nd}	1355.99	1388.08	1354.35	1377.77	1376.38	1310.95	1370.73	1303.05	1343.98	1167.29	1274.32	1336.74	1290.64	1323.30	1050.34	1301.60	1387.14

The smallest MSE is in bold

Table 2 Average PSNR (dB) and SSIM values for the scenarios S_1 , S_2 , S_3 , S_4

Performance	S_1	S_2	S_3	S_4
Average PSNR	27.85	28.64	28.06	29.38
Average SSIM	0.8828	0.8975	0.8860	0.9084

(S_3), SM and DPA are together used to design 15 cluster dictionary pairs. Each cluster dictionary has 600 atoms. SM and DPA are then used together as a model selection criterion according to the proposed algorithm. The same process is repeated in the fourth scenario (S_4) but with perfect model selection. Average PSNR and SSIM values for the four aforementioned scenarios are listed in Table 2 and denoted by S_1 , S_2 , S_3 and S_4 , respectively.

In view of Table 2, it can be concluded that using 15 cluster dictionaries each of 600 atoms defined by SM and DPA is better than using 3 SM cluster dictionaries of 1000 atoms. Perfect model selection results of scenario four (S_4) indicate that there is a significant room for improvement if good

model selection is employed. We now turn to investigating the performance of the proposed algorithm. PSNR and SSIM values of SR outcomes of the aforementioned algorithms are provided in Table 3. It is noted that we have conducted the simulations with source codes provided by the authors of [1], [22] and [23]. Two cases are presented for the proposed algorithm. In the first case, all the cluster dictionary pairs are used. This is denoted by (Prop. 600). In the second case, any LR patch that is classified into C_1 is super-resolved with bicubic interpolation. The second case is denoted by (Prop. 600+Bic.). The proposed algorithm in both cases performs better than the algorithm of Yang et al. [1]. The proposed algorithm has an average PSNR improvements of 0.42 and 0.35 dB over the algorithm of Yang et al. for the first and the second cases, respectively.

In view of Table 3, one notices that the success of the proposed algorithm is particularly valid for images with sharp features such as Text image 3, butterfly and ppt3 images.

It can be seen in Table 3 that the proposed algorithm is competitive with the state-of-the-art algorithms of Peleg et al.

Table 3 PSNR (dB) and SSIM comparisons of bicubic interpolation, the algorithms of Peleg et al. [22], Yang et al. [1] and He et al. [23] and the proposed algorithm, respectively

Image	Bicubic	Yang et al. [1]	Peleg et al. [22]	He et al. [23]	Prop. 600	Prop. 600+Bic.
Barbara	25.35	25.86	25.76	25.84	25.86	25.87
	0.7930	0.8357	0.8359	0.8372	0.8353	0.8350
BDS 198054	24.75	26.85	26.76	26.98	27.19	27.16
	0.8267	0.8816	0.8760	0.8839	0.8835	0.8810
Building image 1	26.41	28.81	29.22	29.21	29.45	29.31
	0.8639	0.9162	0.9219	0.9207	0.9240	0.9196
Build poster	29.95	32.40	32.07	32.77	33.03	32.83
	0.9239	0.9526	0.9478	0.9572	0.9577	0.9532
Butterfly	27.46	31.26	30.96	31.44	32.07	31.93
	0.8985	0.9457	0.9227	0.9463	0.9511	0.9478
Fence	25.05	26.34	26.17	26.22	26.46	26.41
	0.7449	0.8037	0.7967	0.8045	0.8069	0.8031
House	32.76	34.71	34.49	34.84	35.17	35.02
	0.8928	0.9129	0.9136	0.9145	0.9154	0.9138
Kodak 08	24.26	25.51	25.58	25.65	25.78	25.74
	0.7605	0.8198	0.8141	0.8230	0.8256	0.8224
Lena	34.71	36.36	36.59	36.58	36.44	36.39
	0.8507	0.8631	0.8387	0.8647	0.8633	0.8630
ppt3	26.85	29.68	29.71	29.79	30.51	30.44
	0.9372	0.9604	0.9494	0.9621	0.9646	0.9629
Text image 3	17.17	18.43	18.64	18.42	18.81	18.79
	0.7437	0.8143	0.8309	0.8117	0.8297	0.8275
Texture	20.64	22.55	22.97	22.78	23.05	23.05
	0.8272	0.8939	0.9032	0.8997	0.9040	0.9039
Average	26.28	28.23	28.24	28.38	28.65	28.58
	0.8386	0.8833	0.8792	0.8855	0.8884	0.8861

The best PSNR and SSIM are in bold

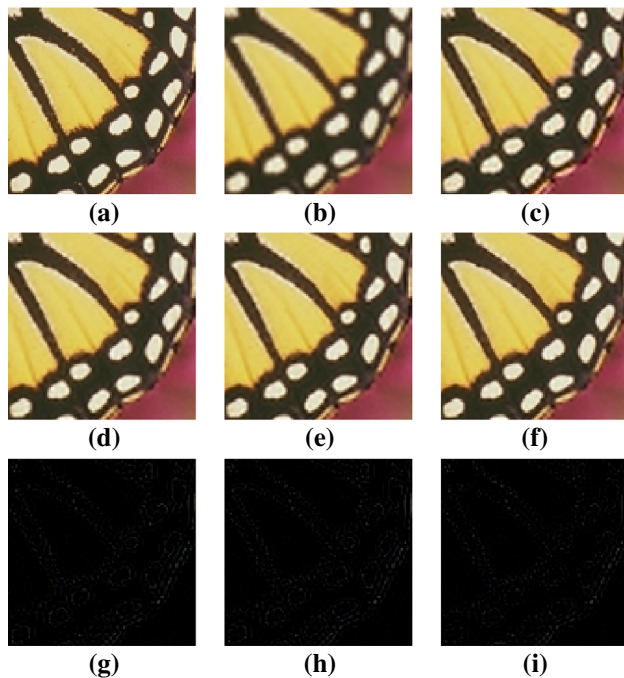


Fig. 4 Results showing a portion of the butterfly image: **a** original image; reconstruction of **b** bicubic interpolation; **c** Peleg et al. [22]; **d** Yang et al. [1]; **e** He et al. [23]; and **f** the proposed algorithm. The *last row* shows the difference between the original image and **g** Yang et al.'s reconstruction, **h** He et al.'s reconstruction and **i** the proposed algorithm's reconstruction

[22] and He et al. [23]. The proposed algorithm has average PSNR improvements of 0.41 and 0.27 dB over the algorithms of Peleg et al. and He et al. respectively. For the case of employing bicubic interpolation for cluster C_1 , the average improvements are 0.34 and 0.20 dB, respectively. SSIM results validate the above observations.

Figure 4 compares a portion of the original butterfly image to its reconstructions obtained with bicubic interpolation, Peleg et al. [22], Yang et al. [1], He et al. [23] and the proposed algorithm. Clearly, the proposed algorithm produces the best reconstruction. This is particularly observed by comparing the curvy line along the butterfly's wing and the patterns on the wing. Figure 4g–i shows, respectively, the differences between the original scene and its reconstructions from Yang et al. [1], He et al. [23] and proposed algorithm. Clearly, the proposed algorithm has the least amount of artifacts.

5 Conclusion

In this paper, a new single-image super-resolution algorithm based on sparse representation over multiple coupled dictionary pairs is proposed. The proposed algorithm clusters image patches based on the magnitude and phase of the gra-

dient operator of image patches. These quantities are used to define sharpness measure and dominant phase angle measures, respectively. Clustering is done in two levels. In the first level, patches are clustered into three clusters based on their sharpness. The second level further clusters patches into several directional sub-clusters based on the dominant phase angle of the gradient operator. For each of the directional sub-clusters, a pair of compact coupled dictionaries are designed. The same two-level clustering paradigm is applied on each LR patch during the reconstruction stage to determine the most appropriate cluster dictionary pair. Sparse coding coefficients of the LR patch over the cluster LR dictionary are calculated. Then, a HR patch estimate is obtained by imposing the sparse coding of the LR patch on the HR dictionary of the same cluster. In this setting, the sharpness and the directional structure of the patch are used together with sparsity as priors to further regularize the super-resolution problem. The designed cluster dictionaries are shown to inherit the sharpness and directional natures of their respective clusters. Tests conducted over natural images show that the proposed algorithm is competitive with the state-of-the-art super-resolution algorithms. The usage of SM- and DPA-based cluster dictionaries contributes to an average PSNR improvement of 0.42 dB over the case of using a single dictionary pair. The average improvements over Peleg et al.'s [22] and He et al.'s [23] algorithms are 0.41 and 0.27 dB, respectively. The improvement in PSNR is particularly significant for images with sharp directional features. SSIM and visual comparison results come inline with the PSNR results.

References

1. Yang, J., Wright, J., Huang, T., Ma, Y.: Image super-resolution via sparse representation. *IEEE Trans. Image Process.* **19**, 2861–2873 (2010)
2. Tian, J., Kai-Kuang, M.: A survey on super-resolution imaging. *SIViP* **5**(3), 329–342 (2011)
3. Yang, J., Wright, J., Huang, T., Ma, Y.: Image super-resolution as sparse representation of raw image patches. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8 (2008)
4. Mairal, J., Bach, F., Ponce, J., Sapiro, G.: Online learning for matrix factorization and sparse coding. *J. Mach. Learn. Res.* **11**, 19–60 (2010)
5. Thanou, D., Shuman, D.I., Frossard, P.: Parametric dictionary learning for graph signals, In: *Proceedings of the IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 487–490 (2013)
6. Mallat, S., Yu, G.: Super-resolution with sparse mixing estimators. *IEEE Trans. Image Process.* **19**(11), 2889–2900 (2010)
7. Dong, W., Zhang, L., Shi, G., Wu, X.: Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. *IEEE Trans. Image Process.* **20**(7), 1838–1857 (2011)
8. Feng, J., Song, L., Yang, X., Zhang, W.: Learning dictionary via subspace segmentation for sparse representation, In: *Proceedings*

- of the IEEE Conference on Computer Vision and Pattern Recognition (ICIP), pp. 1245–1248 (2011)
9. Yu, G., Sapiro, G., Mallat, S.: Image modeling and enhancement via structured sparse model selection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (ICIP), pp. 1641–1644 (2010)
 10. Yang, S., Wang, M., Chen, Y., Sun, Y.: Single-image super-resolution reconstruction via learned geometric dictionaries and clustered sparse coding. *IEEE Trans. Image Process* **21**(9), 4016–4028 (2012)
 11. Kumar, J., Chen, F., Doermann, D.: Sharpness estimation for document and scene images. In: 21st International Conference on Pattern Recognition (ICPR), pp. 3292–3295 (2012)
 12. Sun, J., Xu, Z., Shum, H.Y.: Gradient profile prior and its applications in image super-resolution and enhancement. *IEEE Trans. Image Process* **20**(6), 1529–1542 (2011)
 13. Zuo, W., Zhang, L., Song, C., Zhang, D., Gao, H.: Gradient histogram estimation and preservation for texture enhanced image denoising. *IEEE Trans. Image Process* **23**(6), 2459–2472 (2014)
 14. Yeganli, F., Nazzal, M., Unal, M., Ozkaramanli, H.: Image super-resolution via sparse representation over multiple learned dictionaries based on edge sharpness. *SIViP* (2015). doi:10.1007/s11760-015-0771-7
 15. Dollar, P., Belongie, S., Perona, P.: The fastest pedestrian detector in the west. *BMVC* **2**(3), 7 (2010)
 16. Dalal, N., Triggs, B.: Histogram of oriented gradient for human detection. In: *CVPR* (2005)
 17. Lin, L., Liu, X., Peng, S., Chao, H., Wang, Y., Jiang, B.: Object categorization with sketch representation and generalized samples. *Pattern Recogn.* **45**(10), 3648–3660 (2012)
 18. Peissig, J.J., Young, M.E., Wasserman, E.A., Biederman, I.: The role of edges in object recognition by pigeons. *Perception* **34**(11), 1353–1374 (2005)
 19. Zitnick, C.L., Parikh, D.: The role of image understanding in contour detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 622–629 (2012)
 20. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*, 2nd edn. Prentice-Hall Inc, Englewood Cliffs (2002)
 21. Sun, J., Xu, Z., Shum, H.: Image super-resolution using gradient profile prior. In: Proceedings of IEEE Conference on Computer Vision Pattern Recognition, pp. 1–8 (2008)
 22. Peleg, T., Elad, M.: A statistical prediction model based on sparse representations for single image super-resolution. *IEEE Trans. Image Process.* **23**(6), 2569–2582 (2014)
 23. He, L., Qi, H., Zaretski, R.: Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 345–352 (2013)
 24. Elad, M.: *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer, Berlin (2010)
 25. Elad, M.: Personal webpage. <http://www.cs.technion.ac.il/~elad/software/>. Accessed 1 June 2014
 26. Kodak Lossless True Color Image Suite. <http://r0k.us/graphics/kodak/>. Accessed 1 June 2014
 27. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 416–423 (2001)
 28. Dong, W.: Personal webpage. http://see.xidian.edu.cn/faculty/wsdong/wsdong_downloads.htm. Accessed 1 June 2014
 29. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)