2016 23rd International Conference on Pattern Recognition (ICPR)
Cancún Center, Cancún, México, December 4-8, 2016

Coupled Multiple Dictionary Learning Based on Edge Sharpness for Single-Image Super-Resolution

Junaid Ahmed
Sukkur Institute of Business Administration
Sukkur, Pakistan
Email: junaid_ahmedtl@hotmail.com

Reinhard Klette
Auckland University of Technology
Auckland, New Zealand
Email: rklette@aut.ac.nz

***Abstract*—In this article a new strategy for single-image super-resolution is proposed. A selective sparse coding strategy based on patch sharpness is assumed to be invariant for patch resolution. This sharpness criterion is used at training stage to classify image patches into different clusters. It is suggested that the use of coupled dictionary learning, with a mapping function can improve the representation quality. By this strategy clustered dictionaries are designed along with a mapping function for each cluster which can provide the coupling link between low-resolution and high-resolution image patches. During the reconstruction, image patch sharpness is used as a criterion for the selection of a clustered dictionary and the mapping function. The high-resolution patches are recovered by high-resolution cluster dictionary atoms and the mapping function with sparse representation coefficients from low resolution patches. The algorithm is tested over a set of benchmark images from different data sets. Peak-signal-to-noise ratio and structural-similarity-index measures indicate that the given algorithm is competitive in general with existing baseline algorithms. The proposed algorithm performs better for images with high-frequency components.**

I. Introduction

Currently, sparse representations are extensively used for solving problems in signal and image processing. This representation is also widely used as a regularizer for the *super-resolution* (SR) framework. Consider a signal $\mathbf{x} \in \mathbb{R}^n$ as a vector and $\mathbf{D} \in \mathbb{R}^{n \times k}$, is the associated dictionary where $n \geq 1$ represents the dimension, and $k \geq 1$ represents the columns of $\mathbf{D}$ also called the atoms, the sparse representation of such a scenario can be formulated as follow:

$$\arg \min_{\alpha} ||\mathbf{x} - \alpha \mathbf{D}||_2 \quad s.t. \quad ||\alpha||_0 < S \tag{1}$$

where threshold $S > 0$ models sparsity, $||.||_2$ is the Euclidean norm, and $||.||_0$ is the number of non-zero elements in a vector. The dictionary learning process usually involves the training of data samples by sparse representation to obtain desired dictionary.

Usually it is observed that authors design a set of over-complete dictionaries by introducing the redundancy; though, it has been studied that redundancy can cause degradation in recovery of signals and can induce instabilities [1]. Studies about learning of class-dependent dictionaries follow these observations.

Related work includes Dong et al. [2] where authors applied k-means clustering algorithm to divide the training data at hand into different sets and then applied the dictionary learning to get the compact dictionaries. In [3], Feng et al. use k-subspace clustering and divided the data at hand into different subspaces and then dictionaries were learned from those subspaces in a shared basis manner. More recently, Yu et al. [4] consider the design of structural dictionaries. In [4], considered the orthogonal bases for the dictionary atoms and designed structured dictionaries from those orthogonal bases. Yang et al. [5] propose the use of multiple patch-based clustered dictionaries instead of a single universal dictionary. In this mechanism, the authors studied the geometric properties of the image patches, and patches were clustered into different clusters depending on their geometric property. Dictionaries were obtained from training the image patches from these clusters.

In the reconstruction process of the HR images the existing "sharp" edges, corners, or texture of patches plays an important role. Usually it is observed that patches with such properties are very difficult to reconstruct and many existing SR algorithm fail in doing so. Now these features should be considered for the dictionary learning. Gradient magnitudes can be considered as one of the characterizing parameter for image patches with such features [6]. For an early use of gradient magnitudes, see Rudin et al. [7]. Recently the invariance of the gradient histogram with respect to scale variations (e.g. due to blurring and down-sampling) is applied as a prior for inverse problems; see Sun et al. [8], [9] who define a scale-invariant gradient-profile prior (as a generic image prior) for super-resolution.

The approximate scale-invariant feature introduced as (edge) *sharpness measure* (SM) by Yeganli et al. [10] is proposed to be used as a clustering criterion for classifying image patches into clusters. For each cluster, a pair of compact *high-resolution* (HR) and *low-resolution* (LR) dictionaries is learned along with their mapping functions. In the reconstruction stage, each LR patch is tested with respect to this SM-based clustering criterion. Then, a proper dictionary is selected along with a mapping matrix. Sparse coefficients are calculated for that patch using the selected LR dictionary and mapping matrix. Then, HR patches are reconstructed by using the sparse representation along with the corresponding HR dictionary and mapping matrix. This clustering mechanism, along with the mapping function paradigm, allows us to super-resolve patches with high-frequency components.

978-1-5090-4846-5/16/$31.00 ©2016 IEEE

Simulation results are reported by using the *peak-signal-to-noise ratios* (PSNR) and *structural-similarity-index-measures* (SSIM) [11]. The results indicate that the proposed algorithm is on par with the existing state-of-the-art algorithms and shows improvement in its performance over comparison parameters.

The rest of the paper is organized as follows. Single-image super-resolution via sparse representations is described in Section II. Section III details the proposed SR strategy. Simulation results are shown in Section IV. Section V concludes.

II. Single-Image Super-Resolution via Sparse Representation

Achieving single-image super-resolution is a type of a problem that is ill-posed. Researchers tried to regularize the solution process. Recently, authors proposed a very effective method, called *sparsity*, for regularization. Sparsity has a very nice property of scale invariance (to some extent) due to resolution blur, as described by Yang et al. [12], [13]. Using sparsity as a regularizer, one can find HR images from LR images using the scale invariance of sparse coefficients.

Let $\mathbf{x}_H$ be the HR signal vector extracted from an HR image in the form of a 2D patch, then vectorized into column form. Let $\mathbf{D}_H$ be the corresponding HR dictionary whose columns represent atoms. We can represent this signal vector $\mathbf{x}_H$ by using sparse representations as $\mathbf{x}_H \approx \mathbf{D}_H \alpha_H$, where α_H is a sparse coefficient matrix with only very few non-zero elements.

Let $\mathbf{x}_L$ be the corresponding LR signal vector extracted in the same manner but by blurring and down-sampling the HR image. We can also write the sparse representation of this vector LR signal as $\mathbf{x}_L \approx \mathbf{D}_L \alpha_L$, where $\mathbf{D}_L$ represents the dictionary for the LR signal vector, and α_L represents the sparse coefficient matrix for the LR signal vector.

The relationship between these HR and LR signal vectors is modeled by down-sampling and blurring. This effect is represented by operator Ψ. If such an operation of blurring and down-sampling is performed on HR signals $\mathbf{x}_H$ and the LR signals are obtained $\mathbf{x}_L$, then this can be expressed as:

$$\mathbf{x}_L \approx \Psi \mathbf{x}_H \tag{2}$$

The LR and HR dictionaries are represented as $\mathbf{D}_L$ and $\mathbf{D}_H$, respectively. Also the scale invariance of the sparse coefficients of the LR and HR signals is assumed. Then it follows that these dictionaries are also related by the same operator,

$$\mathbf{D}_L \approx \Psi \mathbf{D}_H \tag{3}$$

See [12]. It follows that

$$\mathbf{x}_L \approx \Psi \mathbf{x}_H \approx \Psi \mathbf{D}_H \alpha_H \approx \mathbf{D}_L \alpha_L \tag{4}$$

From (4) we conclude that $\alpha_H \approx \alpha_L$.

This is the background for a key idea for solving the problem of single-image super-resolution by sparse representation: If we have at hand $\mathbf{x}_L$, $\mathbf{D}_L$, and $\mathbf{D}_H$, then one can calculate α_L by using some vector selection algorithm. Finally the HR patch $\mathbf{x}_H$ can be obtained as:

$$\mathbf{x}_H \approx \mathbf{D}_H \alpha_L \tag{5}$$

III. Proposed Super-Resolution Algorithm

The sharpness measure (SM) $\mathcal{M}$ is defined as an approximate scale-invariant operator for image patches. It is then used for clustering and classification of image patches. Used images are extracted from a set of natural images. (We will map into seven clusters in our experiments.) The derivation of the SM follows from a formula studied in Yeganli et al. [10]:

$$\mathcal{M}(k_1, k_2) = \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \sqrt{|\mathbf{x}_i^h| + |\mathbf{x}_j^v|} \tag{6}$$

where $\mathbf{x}^h$ and $\mathbf{x}^v$ are the directional gradients; h stands for horizontal and v for vertical. Parameters k_1 and k_2 are the dimensions of the image patches.

We apply this approximately scale-invariant measure for clustering. Based on this criterion, seven training data sets are created, with each data set having 5×5 patch dimensions. We sample $40\,000$ of those patches; in vectorized form they represent the data set used for experimental validation.

A. Training Phase

In the training stage, initially a set of natural images is considered for patch extraction. Patches are sampled from this HR data set; as reference to the HR data set, see [13].

From this HR set, each HR image is filtered with a bicubic technique and down-sampled by factor 2 to obtain the corresponding LR image. Such an LR image is then up-sampled by the same technique to obtain a *mid resolution* (MR) image, which has the same resolution as the HR image for easier patch extraction.

Next, patches are sampled from each HR image, and features using gradient filters are extracted from MR images, as already done in [10], [12].

Next, the SM value is calculated for each patch and feature, and, depending on this SM value, patches are classified into seven clusters. Corresponding HR and MR patches, having the same SM value range, are considered as being pairs and classified into the same cluster. In this way, we develop HR and LR training sets. Next, a coupled dictionary learning problem is formulated and solved to obtain the clustered dictionary pairs and their mapping functions.

Let $\mathbf{X}$ and $\mathbf{Y}$ be the HR and LR training sets. We minimize (approximately) the following energy function; by solving this task we obtain the corresponding cluster dictionary for each cluster, along with the needed mapping functions [14]:

$$\begin{aligned} &\min\{\mathbf{D}_\mathbf{x}, \mathbf{D}_\mathbf{y}, f(\cdot)\} E_{\text{data}}(\mathbf{D}_\mathbf{x}, \mathbf{X}) + E_{\text{data}}(\mathbf{D}_\mathbf{y}, \mathbf{Y}) \\ &+ \gamma E_{\text{map}}(f(\alpha_\mathbf{x}), \alpha_\mathbf{y}) + \lambda E_{\text{reg}}(\alpha_\mathbf{x}, \alpha_\mathbf{y}, f(\cdot), \mathbf{D}_\mathbf{x}, \mathbf{D}_\mathbf{y}) \end{aligned} \tag{7}$$

where $E_{\text{data}}(\cdot, \cdot)$ is the data fidelity term, $E_{\text{map}}(\cdot, \cdot)$ is the mapping fidelity, and E_{reg} is the regularizer. The coupling between the sparse coefficients of HR and LR data over

dictionaries is related by the mapping function $f(\cdot)$. The HR and LR dictionaries are optimized concurrently with the mapping function.

The problem in (7) can be converted into a ridge regression and dictionary learning problem considering the mapping to be linear function as:

$$\begin{aligned}\min\{\mathbf{D_x}, \mathbf{D_y}, f(\cdot)\}\|X - \mathbf{D_x}\alpha_\mathbf{x}\|_F^2 + \|Y - \mathbf{D_y}\alpha_\mathbf{y}\| \\ \gamma\|\alpha_\mathbf{y} - \mathbf{M}\alpha_\mathbf{x}\|_F^2 + \lambda_\mathbf{x}\|\alpha_\mathbf{x}\|_1 + \lambda_\mathbf{y}\|\alpha_\mathbf{y}\|_1 + \lambda_m\|\mathbf{M}\|_F^2 \\ s.t.\ \|\mathbf{D}_{x,i}\|_{l_2} \leq 1 \wedge \|\mathbf{D}_{y,i}\|_{l_2} \leq 1\ ,\ \text{for all}\ i \end{aligned} \quad (8)$$

where γ, $\lambda_\mathbf{x}$, λ_m, and $\lambda_\mathbf{y}$ represent the regularization terms for the optimum performance, and $\mathbf{D}_{\mathbf{x},i}$ and $\mathbf{D}_{\mathbf{y},i}$ represent the atoms of the dictionaries $\mathbf{D_x}$ and $\mathbf{D_y}$.

The problem formulated by (8) can be solved by optimizing one parameter at a time while considering the others as being constant. As the mapping function (matrix) $\mathbf{M}$ is linear, bidirectional transforms are learned from $\alpha_\mathbf{x}$ to $\alpha_\mathbf{y}$, and vice versa.

After initializing matrix $\mathbf{M}$ and dictionary $\mathbf{D}$, one can find the sparse coefficients α by applying

$$\begin{aligned}\min\{\alpha_\mathbf{x}\}\|X - \mathbf{D_x}\alpha_\mathbf{x}\|_F^2 + \gamma\|\alpha_\mathbf{y} - \mathbf{M_x}\alpha_\mathbf{x}\|_F^2 + \lambda_\mathbf{x}\|\alpha_\mathbf{x}\|_1 \\ \min\{\alpha_\mathbf{y}\}\|Y - \mathbf{D_y}\alpha_\mathbf{y}\|_F^2 + \gamma\|\alpha_\mathbf{x} - \mathbf{M_y}\alpha_\mathbf{y}\|_F^2 + \lambda_\mathbf{y}\|\alpha_\mathbf{y}\|_1 \end{aligned} \quad (9)$$

The problem in (9) can easily be solved by applying an l_1 norm minimization algorithm such as LARS [15].

Now for the dictionary update stage using the current sparse coefficients the following problem is solved:

$$\begin{aligned}\min\{\mathbf{D_x}, \mathbf{D_y}\}\|\mathbf{X} - \mathbf{D_x}\alpha_\mathbf{x}\|_F^2 + \|\mathbf{Y} - \mathbf{D_y}\alpha_\mathbf{y}\|_F^2 \\ s.t.\ \text{for all}\ i, \|\mathbf{D}_{x,i}\|_{l_2} \leq 1 \wedge \|\mathbf{D}_{y,i}\|_{l_2} \leq 1 \end{aligned} \quad (10)$$

Now this is particularly called a *quadratically constrained quadratic program* (QCQP) problem. It can be solved as outlined in [13]. Finally, by keeping the dictionary and the sparse coefficients fixed, matrix $\mathbf{M}$ can be updated as follows:

$$\min\{\mathbf{M}\}\|\alpha_\mathbf{y} - \mathbf{M}\alpha_\mathbf{x}\|_F^2 + (\lambda_m/\gamma)\|\mathbf{M}\|_F^2 \quad (11)$$

The problem in Equation (11) is called the ridge regression problem; its solution can be given as:

$$\mathbf{M} = \alpha_\mathbf{y}\alpha_\mathbf{x}^T(\alpha_\mathbf{x}\alpha_\mathbf{x}^T + (\lambda_m/\gamma)\cdot I)^{-1} \quad (12)$$

where $\mathbf{I}$ represents the identity matrix.

By this strategy we develop a set of clustered dictionaries, for each SM-based cluster, along with their mapping function (matrix). The proposed training algorithm is summarized in Fig. 1.

B. Reconstruction Phase

At the image recovery stage, a given LR image is first upward converted into MR level by a bicubic interpolation technique. This is done for matching the size of the HR and the (now transformed) LR image. Patches and features are then extracted from this up-converted image applying a full-overlap selection scheme.

This is followed by the selective sparse coding step. We need to identify which dictionary pair and mapping function is to be used for the sparse recovery of a patch.

This is done in the same way as we clustered the patches by using the approximate scale invariant SM. Next, the sparse coefficients of this patch are calculated using the LR dictionary and the mapping matrix of that cluster. The next step is to reconstruct the HR patch. This is done by using the scale invariance property of the sparse coefficients, and by using the sparse coefficients of the LR versions along with the HR dictionary for approximating the HR patches.

Finally we reshape these approximated HR vector patches into the 2-dimensional form, and, as we know patches were extracted with full overlap, we employ the overlap-addition method by [12] and arrive at our approximate HR image. The reconstruction process is summarized in Fig. 2.

This proposed strategy of clustering, based on patch sharpness and fully coupled dictionary learning (by mapping functions), helps to improve the overall quality of the reconstructed HR image. The mapping function serves for coupling the LR and HR sparse coefficients. It helps to reconstruct better quality HR patches from their LR versions. The proposed SM-based clustering serves in particular the purpose of recovering high-frequency components of HR images.

IV. Experiments and Evaluations

We present simulation parameters and results for evaluating the performance of our proposed algorithm. We compare with the algorithm of Yang et al. [12], which is considered to be the baseline algorithm, the algorithm by Yeganli et al. [10], and bicubic interpolation. Experiments are performed on a set of benchmark images chosen from various data sets including Kodak and other data sets publicly available by [16], [17].

Algorithm 1 The Training Phase

Input: HR Training Image Set.

- Extract patches from HR images and features from LR.
- Based on the SM value of HR patch and LR feature classify them into the seven clusters.
- For each cluster concatenate the patches and features into adjacent pairs and formulate the HR and LR training sets.

For **X (HR)** and **Y (LR)** training sets of seven clusters:

- Initialize dictionary pairs $\mathbf{D_x}$ and $\mathbf{D_y}$ for each cluster.
- Initialize mapping functions $\mathbf{M_x}$ and $\mathbf{M_y}$ for each cluster.

Now take one cluster at a time and do:

- Fix other variables, update $\boldsymbol{\alpha_x}$ and $\boldsymbol{\alpha_y}$ by Eq. 9.
- Fix other variables, update $\mathbf{D_x}$ and $\mathbf{D_y}$ by Eq. 10.
- Fix other variables, update $\mathbf{M_x}$ and $\mathbf{M_y}$ by Eq. 11.

Output: $\mathbf{D_x}$, $\mathbf{D_y}$,$\mathbf{M_x}$ and $\mathbf{M_y}$ for each cluster.

Fig. 1. Proposed dictionary learning algorithm

Algorithm 2 The Reconstruction Phase

Input: LR Test Image and Cluster Dictionary Pairs with Mapping Functions.

- Up-sample the LR test image to get the MR image.
- Extract patches with full overlap and apply feature extraction filters on MR images.
- Concatenate the overlapping MR features in column vectors.

Now **for** each patch **do:**

- Find its SM value and locate the dictionary and mapping function.
- Given the cluster dictionary and patch, find the sparse representation of that patch.
- Given the sparse representation using the mapping function and the corresponding HR dictionary calculate the HR patch estimate.

End for:

- Get the HR image estimate by merging the patches by overlap-add method.

Output: A HR image estimate.

Fig. 2. Proposed reconstruction algorithm

A. Quantitative Experimentation

First, the HR images are down-sampled by a bicubic technique to the obtain the LR versions, and then these LR images are reconstructed by the three considered algorithms and the bicubic technique to their original sizes. PSNR and SSIM measures, as given in [11], [12], are used along with sharpness and contrast measures defined by Liu et al. [18], for quantitative performance evaluation.

The PSNR measure for a reconstructed image is calculated as follows:

$$\mathcal{M}_{\mathrm{PSNR}}(\mathbf{x},\hat{\mathbf{x}}) = 10\log_{10}\frac{255^2}{\mathcal{E}_{\mathrm{MSE}}(\mathbf{x},\hat{\mathbf{x}})} \tag{13}$$

where $\mathbf{x}$ is the original HR image having size of $M \times N$, and $\hat{\mathbf{x}}$ is the estimation, and $\mathcal{E}_{\mathrm{MSE}}(\mathbf{x},\hat{\mathbf{x}})$ is the *mean square error* (MSE) given for $\mathbf{x}$ and $\hat{\mathbf{x}}$ as follows:

$$\mathcal{E}_{\mathrm{MSE}}(\mathbf{x},\hat{\mathbf{x}}) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(\mathbf{x}_{ij}-\hat{\mathbf{x}}_{ij})^2 \tag{14}$$

We also included SSIM [11] as a perceptual quality metric, which is more compatible with human image quality perception than the PSNR measure. The sharpness and contrast measures, as introduced by Liu et al. [18], are at first calculated as $s(i,j)$ and $c(i,j)$, respectively, for each pixel position (i,j) and then averaged for the whole image.

Regarding $s(i,j)$ and $c(i,j)$, consider an image I and $A(i,j)$ as being the 8-adjacent pixels "around" (i,j) (not including (i,j)); then

$$s(i,j) = \|I(i,j) - \mu_{A(i,j)}\|_1 \tag{15}$$

where $s(i,j)$ is the sharpness value of image I at (i,j), and $\mu_{A(i,j)}$ the mean value of I at pixel locations in $A(i,j)$. For the contrast, let

$$c(i,j) = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}\|I(i,j) - I(x,y)\|_1 \tag{16}$$

where $c(i,j)$ is the contrast value of image I at (i,j).

Fig. 3. Test images *from left to right, and top to bottom*: `Peppers`, `Rocio`, `Starfish`, `Yan`, `AnnieYukiTim`, `Barbara`, `BooksCIMAT`, `Butterfly`, `Fence`, `ForbidenCity`, `HowMany`, `Kodak-05`, `Kodak-08`, `Michoacan`, `MissionBay`, and `NuRegions`

Figure 3 shows a set of test images used for experimental tests. This set of images is prepared from a predefined set of single images [16], [17] available online. Some of the images are benchmarks and used widely by the researchers in the field of pattern recognition and image processing such as `Butterfly`, `Peppers`, `Barbara`, `Fence`, and `Starfish`.

Table I lists the PSNR and SSIM [11] values, and the sharpness and contrast values for the conducted experiments, all rounded to two decimal places. In a few cases, maxima (shown in bold) have been identified using more than two decimal places.

We used a patch size of 5×5, and the trained dictionaries have a size of 25×256. Dictionaries and mapping functions are learned for each cluster, with $40\,000$ patches sampled for each cluster.

The algorithms of [10] and [12] have also been used with same simulation parameters to avoid any unfair advantage. Also, a single data set of natural images, as used by [12], is applied for dictionary learning for all the algorithms. For bicubic interpolation, we simply used MATLAB's command (`imresize`) with scale factor of two. The algorithm by [10] is chosen for comparison because it uses a similar SM-measure for clustering, and it also employs the coupled dictionary learning of [12].

In comparison with the baseline algorithm of [12], the proposed algorithm uses seven dictionaries for recovering any of the HR patches. But note that for each patch only one dictionary out of the seven is used for its recovery. As the dictionaries are learned off-line they do not incur any additional computational cost. The only additional computation

occurs for the sharpness measure for the LR patch for cluster dictionary selection.

The sharpness and contrast values are for comparing the contrast and sharpness values of reconstructed images with those of the original images. The table shows absolute errors (i.e. the absolute difference in contrast or sharpness from the original value, divided by the original value). Smaller values indicate less deviation from true contrast and sharpness. The table indicates that images reconstructed by the proposed algorithm have less deviation in terms of sharpness from the original value. This corresponds to the observation that the proposed algorithm is well able to recover high-frequency components better than the other algorithms. We have slightly more deviation from the original contrast value compared with the other algorithms.

Table I also shows that the proposed algorithm has an average PSNR improvement of 0.46 dB over the algorithm of Yang et al. [12], and a 0.35 dB improvement over the algorithm of Yeganli et al. [10]; the improvement of the proposed algorithm over the bicubic interpolation is 0.94 dB. These PSNR improvements, documented in the table, are visually evident in the tested images. For example, considering the `Butterfly`, `Rocio`, an `NuRegions` images, the results are self-explanatory.

Regarding the SSIM values, the proposed algorithm shows an improvement in average of 0.0172 over the Yang et al. [12] algorithm, of 0.0058 over the Yeganli et al. [10] algorithm, and of 0.0347 over the bicubic technique.

TABLE I
PSNR (TOP-LEFT), SHARPNESS (TOP-RIGHT), SSIM (BOTTOM-LEFT), AND CONTRAST (BOTTOM-RIGHT) VALUES FOR COMPARISON OF BICUBIC INTERPOLATION, YANG ET AL.'S [12] ALGORITHM, YEGANLI ET AL.'S [10] ALGORITHM, AND THE PROPOSED ALGORITHM

Images	Bicubic		[12]		[10]		Proposed	
AnnieYukiTim	31.42 0.91	0.57 0.32	32.68 0.92	0.39 0.20	32.75 0.92	0.38 0.20	**33.08** **0.92**	**0.36** 0.21
Barbara	25.35 0.79	0.74 0.59	25.7 0.83	0.58 0.44	25.70 0.83	0.57 0.44	**25.83** **0.83**	**0.54** 0.43
BooksCIMAT	28.35 0.90	0.51 0.31	30.42 0.93	0.27 0.12	30.56 0.93	0.26 0.12	**30.86** **0.94**	**0.23** 0.12
Butterfly	27.45 0.90	0.45 0.16	29.90 0.92	0.22 0.03	30.10 0.93	0.21 0.02	**31.36** **0.95**	**0.19** 0.06
Fence	25.04 0.74	0.64 0.53	26.29 0.77	0.45 0.33	26.40 0.80	0.46 0.33	**26.47** **0.80**	**0.44** 0.35
ForbiddenCity	26.92 0.82	0.62 0.44	28.01 0.86	0.41 0.25	28.03 0.86	0.41 0.25	**28.21** **0.87**	**0.39** 0.26
HowMany	31.67 0.93	0.46 0.22	33.65 0.93	0.24 0.08	33.77 0.93	0.24 0.08	**33.87** **0.94**	**0.21** 0.09
Kodak-05	26.97 0.83	0.58 0.39	28.46 0.88	0.36 0.20	28.58 0.88	0.35 0.20	**28.96** **0.89**	**0.33** 0.21
Kodak-08	24.25 0.76	0.66 0.47	25.09 0.80	0.47 0.30	25.19 0.81	0.46 0.29	**25.50** **0.82**	**0.44** 0.31
Michoacan	36.31 0.87	0.57 0.35	27.70 0.89	0.36 0.19	27.82 0.90	0.35 0.19	**28.10** **0.90**	**0.32** 0.20
MissionBay	29.80 0.90	0.52 0.28	31.61 0.92	0.30 0.12	31.69 0.92	0.29 0.12	**32.33** **0.93**	**0.27** 0.14
NuRegions	24.30 0.92	0.41 0.14	26.90 0.85	0.19 0.01	27.06 0.95	0.17 0.01	**27.50** **0.96**	**0.15** 0.02
Peppers	33.05 0.84	0.65 0.39	33.80 0.84	0.50 0.29	33.96 0.84	0.50 0.29	**34.15** **0.84**	**0.48** 0.30
Rocio	36.63 0.96	0.40 0.16	38.89 0.97	0.18 0.06	39.01 0.97	0.18 0.06	**39.22** **0.97**	**0.14** 0.07
Starfish	30.22 0.89	0.51 0.28	31.94 0.92	0.30 0.13	32.04 0.92	0.29 0.13	**32.33** **0.93**	**0.27** 0.14
Yan	29.86 0.89	0.56 0.32	31.35 0.91	0.34 0.15	31.46 0.90	0.34 0.15	**31.95** **0.91**	**0.32** 0.17
Means	29.22 0.87	0.55 0.33	30.15 0.88	0.35 0.18	30.26 0.89	0.34 0.18	30.61 0.90	0.32 0.19

B. Qualitative Experimentation

Here we show zoomed-in versions of three reconstructed images just for a visual comparison. Fig. 4 and Fig. 5 show the zoomed-in original image and the images reconstructed by the algorithms used for comparison.

We zoomed into the images to better visual comparison. Looking at Fig. 4 and Fig. 5 one can see that the reconstruction by the bicubic technique shows a significant amount of blur. The reconstructed images by using the algorithm of Yang et al. [12] or Yeganli et al. [10] are much clearer then those by the bicubic technique. Looking at the zoomed-in fence image, we can clearly see that the reconstruction by the proposed algorithm is much sharper around the edges. In general we can say that the proposed algorithm is able to recover the

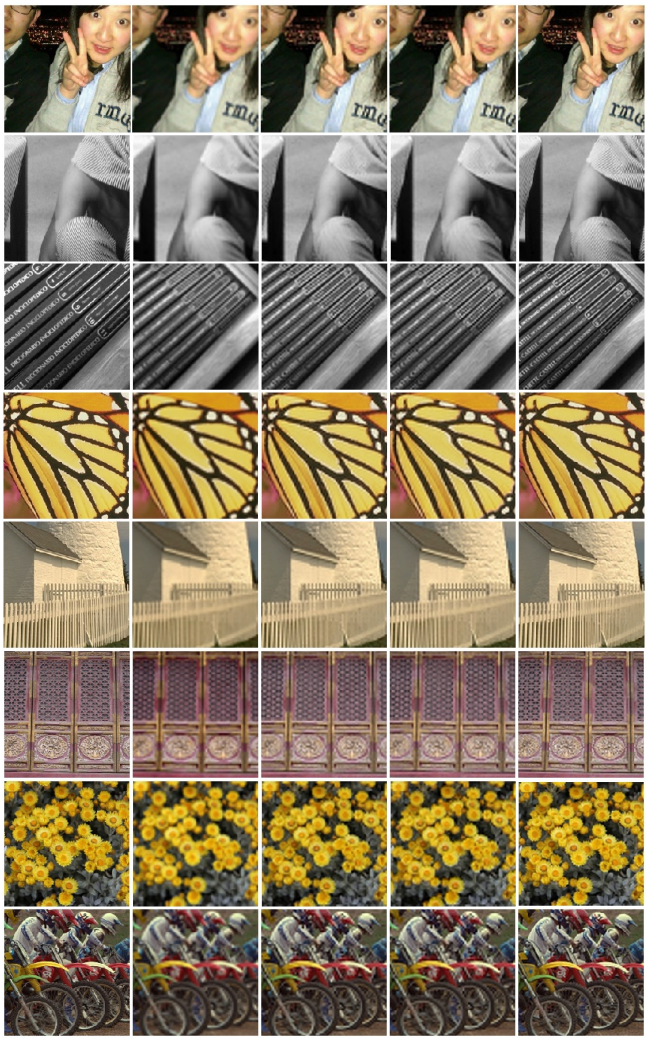

Fig. 4. Visual comparison of results for images `AnnieYukiTim`, `Barbara`, `BooksCIMAT`, `Butterfly`, `Fence`, `ForbiddenCity`, `HowMany`, and `Kodak05`, from left to right: Original image, bicubic interpolation, algorithm of [12], algorithm of [10], and proposed method

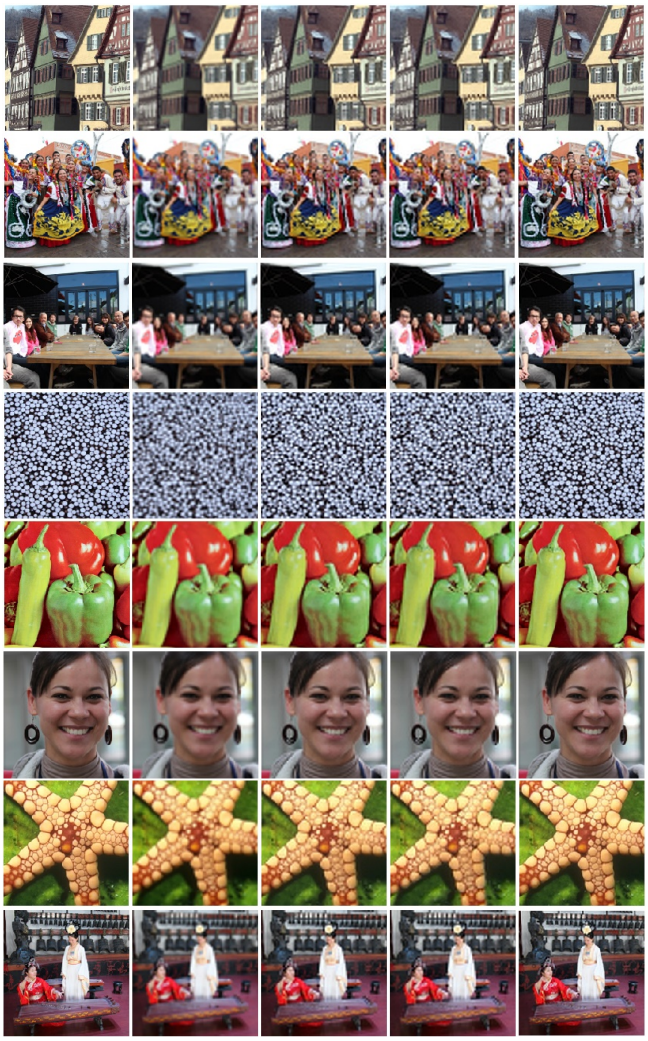

Fig. 5. Visual comparison of results for images `Kodak08`, `Michoacan`, `NuRegions`, `Peppers`, `Rocio`, `Starfish`, and `Yan`, from left to right: Original image, bicubic interpolation, algorithm of [12], algorithm of [10], and proposed method

sharper patches more efficiently than the baseline algorithm used for comparison.

V. Conclusions

We used a coupled dictionary and mapping learning on a set of clusters classified by SM values of the HR and LR patches. The sharpness measure is here used as a classification criterion for the patches. We tested on patches of size 5×5 with a dictionary of size 25×256 to ensure computational efficiency.

Experimental results show that the proposed idea of selective sparse coding over sharpness-clustered dictionaries is useful, combined with the use of a coupled dictionary and mapping function learning. This altogether strengthens the concept of super-resolution recovery by sparse representation. The coupled dictionary and mapping learning further couples the sparse representation coefficients of the HR and LR images. This helps for improving super-resolution reconstruction. Compared to bicubic interpolation, the proposed algorithm gives a 0.94 dB improvement when tested over a large set of benchmark images. The proposed algorithm provides a 0.46 dB improvement over the baseline algorithm of Yang et al. [12], and of 0.35 dB over the algorithm of Yeganli et al. [10]. Visual evaluations also correspond to those quantitative results.

It is recommended to choose the best dictionary for reconstruction (from the set of selected dictionaries) based on the sparse representation error.

References

[1] S. Mallat and G. Yu, Super-resolution with sparse mixing estimators, *IEEE Trans. Image Processing*, 19:2889–2900, 2010.

[2] W. Dong, L. Zhang, G. Shi, and X. Wu, Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization, *IEEE Trans. Image Processing*, 20:1838–1857, 2011.

[3] J. Feng, L. Song, X. Yang, and W. Zhang, Learning dictionary via subspace segmentation for sparse representation, in Proc. *IEEE Int. Conf. Image Processing*. pp. 1245–1248, 2011.

[4] G. Yu, G. Sapiro, and S. Mallat, Image modeling and enhancement via structured sparse model selection, in Proc. *IEEE Int. Conf. Image Processing*. pp. 1641–1644, 2010.

[5] S. Yang, M. Wang, Y. Chen, and Y. Sun, Single-image super-resolution reconstruction via learned geometric dictionaries and clustered sparse coding, *IEEE Trans. Image Processing*, 21:4016–4028, 2012.

[6] J. Kumar, F. Chen, and D. Doermann, Sharpness estimation for document and scene images, in Proc. *Int. Conf. Pattern Recognition*, pp. 3292–3295, 2012.

[7] L. I. Rudin, S. Osher, and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenomena*, 60:259–268, 1992.

[8] J. Sun, J. Sun, Z. Xu, and H.-Y. Shum, Image super-resolution using gradient profile prior, in Proc. *IEEE Conf. Computer Vision Pattern Recognition*, pp. 1–8, 2008.

[9] J. Sun, Z. Xu, and H.-Y. Shum, Gradient profile prior and its applications in image super-resolution and enhancement, *IEEE Trans. Image Processing*, 20:1529–1542, 2011.

[10] F. Yeganli, M. Nazzal, M. Unal, and H. Ozkaramanli, Image super-resolution via sparse representation over multiple learned dictionaries based on edge sharpness, *Signal Image Video Processing*, 10:535–542, 2016.

[11] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, *IEEE Trans. Image Processing*, 13:600–612, 2004.

[12] J. Yang, Z. Wang, Z. Lin, S. Cohen, and T. Huang, Coupled dictionary training for image super-resolution, *IEEE Trans. Image Processing*, 21:3467–3478, 2012.

[13] J. Yang, J. Wright, T. S. Huang, and Y. Ma, Image super-resolution via sparse representation, *IEEE Trans. Image Processing*, 19:2861–2873, 2010.

[14] S. Wang, L. Zhang, Y. Liang, and Q. Pan, Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis, in Proc. *IEEE Conf. Computer Vision Pattern Recognition*, pp. 2216–2223, 2012.

[15] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, Least angle regression, *Annals Statistics*, 32:407–499, 2004.

[16] R. Franzen. Kodak lossless true color image suite, online: r0k.us/graphics/kodak/index.html [accessed 20-January-2016].

[17] R. Klette. *Concise Computer Vision*, Springer, London 2014. Single images, online: ccv.wordpress.fos.auckland.ac.nz/data/single-images/ [accessed 20-January-2016].

[18] D. Liu and R. Klette, Sharpness and contrast measures on videos, in Proc. *Image Vision Computing New Zealand*, IEEE online, 2015.