IMAGE SUPER RESOLUTION USING SALIENCY-MODULATED CONTEXT-AWARE SPARSE DECOMPOSITION

Wei Bai, Saboya Yang, Jiaying Liu*, Jie Ren and Zongming Guo

Institute of Computer Science and Technology, Peking University, Beijing, P. R. China 100871

ABSTRACT

This paper presents a novel saliency-modulated sparse representation algorithm for image super resolution. In images, regions salient to human eyes appear to be more organized and structured. This property is utilized in both the dictionary learning and the sparse coding process to capture more structural details for the reconstructed image. Apart from a general dictionary, example patches from the salient regions are extracted to train a salient dictionary. We also incorporate context-aware sparse decomposition to model dependencies between dictionary atoms of adjacent patches, especially in the salient regions. Experiments show the proposed method outperforms state-of-the-art methods with the highest PSNR gain. Subjective results demonstrate the proposed method reduces artifacts and preserves more details.

Index Terms— Super resolution, sparse representation, saliency, context-aware

1. INTRODUCTION

Sparse representation of signals on over-complete dictionaries is a rapidly evolving field. The basic model suggests that natural signals can be compactly expressed as a linear combination of prespecified atom signals, where the linear coefficients are sparse (*i.e.*, most of them zeros). Formally, let $x \in \mathbb{R}^n$ be a column signal, and $D \in \mathbb{R}^{n \times m}$ be a dictionary, the sparsity assumption is described by the following sparse approximation problem:

$$x \approx D\gamma, \ \ s.t. \|\gamma\|_0 \leq \epsilon. \tag{1}$$

where γ is the sparse representation of x, ϵ is a predefined threshold. The l_0-norm $\| \cdot \|_0$ counts the nonzero entries of a vector, claiming the sparsity of x. Though l_0-norm optimization is a NP-hard problem, there are various ways to solve it [1, 2].

Sparse representation-based super-resolution (SR) techniques are extensively studied in recent years. They attempt to capture the co-occurrence prior between low-resolution

(LR) and high-resolution (HR) image patches. Yang *et al.* [3] used a coupled dictionary learning model for image super-resolution. They assumed that there exist coupled dictionaries of HR and LR images, which have the same sparse representation for each pair of HR and LR patches. After learning the coupled dictionary pair, the HR patch is reconstructed on HR dictionary with sparse coefficients coded by LR image patch over the LR dictionary. In this typical framework of sparse representation-based SR method, the dictionary is determined on a general training set and the prior model to constrain the restoration problem is the sparsity of each local patch.

Many dictionary learning methods aim at learning a universal dictionary on a general training set to represent various image structures [4]. However, for complex natural images, sparse decomposition over a highly redundant dictionary is potentially unstable and tends to generate visual artifacts. In other words, universal dictionaries are not adaptive to local image properties. Therefore, it is reasonable to improve the dictionary learning model for more adaptive dictionaries. Fortunately, the rapid development of social network provides us with large amount of similar images describing the same scene. This means similar images of the LR image can be gathered to train an adaptive dictionary.

Moreover, inspired by the work in [5], we consider introducing the saliency property of images to further improve the adaptiveness of the dictionary. Saliency refers to elements of a visual scene that are likely to attract the attention of human observers [6]. More generally, regions salient to human eyes tend to be highly structured because human visual system is attracted to organized structures for the ease of recognition. Sadaka *et al.* also suggested that due to human visual attention, attended regions are processed at high visual acuity, hence details in these regions should be reconstructed with higher accuracy than those in non-attended areas. Thus when training dictionaries, we specially use samples from salient regions. The fact that salient regions of similar images probably have similar structures would enhance the adaptiveness and reconstruction ability of dictionaries. When reconstructing the image, the above trained dictionary can be applied to the salient regions in the LR image to generate more visually pleasant results while reducing the overall computation cost.

As to the prior model to recover the HR image, conventional sparse recovery algorithms [7] imposed the sparsity constraint of each independent patches. The local smooth-
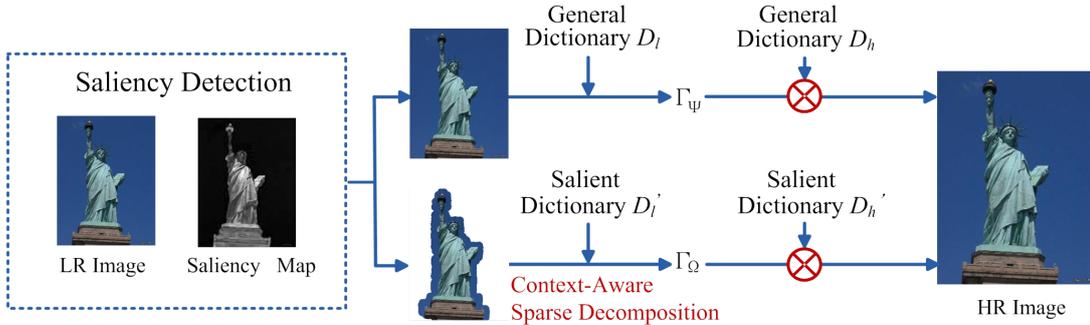
Fig. 1. Flow diagram of the proposed algorithm.

ness is constrained merely by averaging on overlapped regions, which is weak to regularize the image SR problem when the observed LR image loses partial structure information. Correlations between the structural information of the whole patches (not merely the overlapped regions) should be investigated. Thus context-aware sparse decomposition is introduced, which refers to sparsely coding the patches by employing the dependencies between the dictionary atoms used to decompose the patches. Better still, the highly structured property of salient regions makes it a fairly proper scenario to apply context-aware sparse coding.

In this paper, considering the aforementioned two issues, we present a novel saliency-modulated context-aware sparse decomposition method for image super resolution. Similar images are obtained from the Internet by content-based image retrieval to build a specialized database. Then example patches from the salient regions of the database are extracted to train a salient dictionary, which is especially adaptive to local structures. In addition, to better explore the correlations among patches, we apply context-aware sparse decomposition to salient regions based on the observation that salient regions tend to be more structured.

The rest of this paper is organized as follows: Section 2 describes each part of the proposed algorithm in detail. Experimental results are shown in Section 3. Finally, concluding remarks are given in Section 4.

2. SALIENCY-MODULATED CONTEXT-AWARE SPARSE DECOMPOSITION

2.1. Overview

The sparse representation-based SR problem can be formulated as given a low-resolution image, recovering its high-resolution version via the learned coupled dictionaries. $X = \{x_1, x_2, ..., x_t\}$ is a set of training examples (all of them have been reformed to column signals), then the conventional dictionary learning process aims at minimizing the following formulation [3]:

$$D = \arg\min_{D,\Gamma} \|X - D\Gamma\|_2^2 + \lambda\|\Gamma\|_0, \quad s.t.\|D_i\|_2^2 \leq 1, \quad (2)$$

where $\|\Gamma\|_0$ is the sparsity constraint and $\|X - D\Gamma\|_2^2$ is the data fidelity constraint. D_i $(i = 1, 2, ..., K)$ represents the atoms of the dictionary D. This extensively-studied l_0-norm minimization problem can be approximated by greedy algorithms or convex relaxation-based algorithms. A coupled dictionary which includes both the LR and HR dictionary can be trained in the similar way. Once the dictionary is settled, the LR image patch x can be sparsely coded as follows:

$$\hat{\gamma} = \arg\min_{\gamma} \|X - D\gamma\|_2^2 + \lambda\|\gamma\|_0. \quad (3)$$

And the problem of recovering the HR patch $\hat{y}$ turns into multiplying the sparse coefficients γ with the HR dictionary.

Conventional methods randomly choose patches to build training set X, which results in a general dictionary. But for a particular LR image, this dictionary is too general to express certain structural details. We improve the dictionary learning model from two levels to enhance the adaptiveness of learned dictionaries. First, similar images of the LR image are gathered to be candidate training set. Then, we go on to narrow it down to salient parts of images in the training set. The proposed dictionary training procedure is based on two important facts: 1) Similar images contain more useful information than general images that will help compensate for the LR image. 2) Salient regions are highly structured, signals extracted from the salient regions should be closely correlated. So the learned dictionary is especially adaptive to the structure of the salient regions. Since attended salient regions need to be treated with more acuity from the human visual perspective, we apply salient dictionaries only to salient regions. For the less attended non-salient regions, a general dictionary will do.

In the sparse coding phase shown in eq.(3), sparsity of each independent patches is used to regularize the optimization problem. In other words, each patch is sparsely coded independently and overlapped regions are averaged to keep smoothness along boundaries. However, neighboring patches are closely correlated. They may tend to have similar sparse codes. Especially for patches in salient regions, the highly structured property makes them more dependent on each other. Thus we introduce the context-aware sparse decomposition to employ the dependencies between the dictionary atoms used to decompose the patches. Such improvement on

the prior model imposes more constraints on the restoration problem, which will help preserve more structural details.

Based on the characteristics of saliency, the proposed algorithm generates adaptive dictionaries and also sparsely decompose patches in a correlated way. A set of similar images to the LR image are collected in Ψ, and then comes the salient database Ω. Obviously, $\Omega \subseteq \Psi$. Let X_Ψ be a set of patches which are extracted from the whole images in the database, and X_Ω be the patches extracted from the salient regions in the images of the database. Then the general dictionary, $\{D_l, D_h\}$, and the saliency-modulated dictionary, $\{D_l', D_h'\}$, are obtained by training on X_Ψ and X_Ω respectively using the method described in Sec.2.1. Γ_Ψ and Γ_Ω are sparse codes for non-salient and salient regions, as Fig.1 illustrates. Note that, Γ_Ω is obtained by context-aware sparse coding. Therefore, salient regions are reconstructed with more accuracy owing to the context-aware sparse coding process.

Hence, in the proposed scheme, the most important parts are saliency segmentation and modeling the correlation network of local patches. We will elaborate on them in the following sections.

2.2. Salient dictionary learning

The difference between general dictionary and salient dictionary is the choice of training examples. Instead of using examples distributed all over the database, we only choose patches from the salient regions of the images in the database, as Fig.1 shows. Naturally we get a dictionary which is especially adaptive to the structure of the salient regions.

With regard to choosing patches from the salient regions of the images, we have to detect and segment salient regions first. A simple but efficient approach developed in [8] is adopted to tackle this problem. It identifies salient regions as those regions of an image that are visually more conspicuous by virtue of their contrast with respect to surrounding regions. They use a contrast determination filter that operates at various scales to generate saliency maps containing "saliency values" per pixel.

Fig.2 shows the results of the saliency detection and segmentation operation. Fig.2(b) reveals the saliency values of the original image, which is consistent with common sense. On the basis of the saliency map, mean-shift based segmentation is performed to crop out the salient region in Fig.2(d).

After the specific salient signals are chosen, a salient dictionary is learned as in eq.(2). In this work, we use the *SPAMS*[1] Matlab package to train the general and the salient dictionaries.

2.3. Context-aware sparse decomposition in salient regions

Instead of enforcing the compatibility of overlapped regions between neighboring patches, we investigated the context-

(a) Original	(b) Saliency Map
(c) Saliency Segmentation	(d) Salient Region
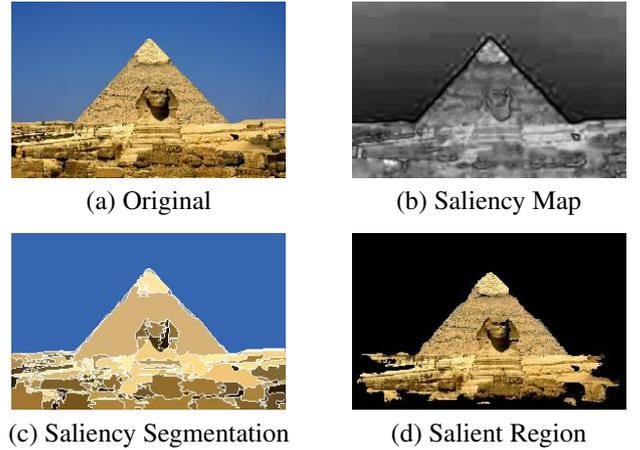
Fig. 2. An example of salient region segmentation.

aware sparse decomposition of patches, which means the correlations between the structural information of the whole patches, not only in the overlapped regions, are explored. Correlations between the structural components of the adjacent patches refer to the dependencies between the dictionary atoms which are used to decompose the patches. As mentioned before, regions that are salient to human eyes tend to be highly structured and probably share similar sparse codes. This provides a reasonable scenario to apply the context-aware sparse decomposition.

Let γ_i be the sparse representation vector of the current patch x_i, and $\gamma_{i \diamond t}, t = 1, 2, ...8$, be the sparse codes of x_i's neighbor patches in 8 directions (see Fig.3, *e.g.*, the patch in dash line stands for direction-1 patch). Denote S_i as sparsity pattern of representation γ_i ($S_i \in \{-1, 1\}^m$), if $\gamma_i(j) \neq 0$ (*i.e.*, the j-th atom of γ_i) then $S_i(j) = 1$, otherwise $S_i(j) = -1$. $S_{i \diamond t}$ represents the sparsity pattern of the adjacent patch in orientation t.
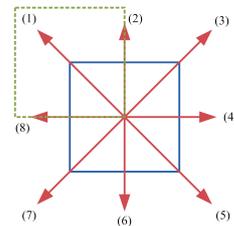
Fig. 3. The local neighborhood system of patch x_i with a spatial configuration of eight different orientations.

Given all the orientated neighboring sparsity patterns $\{S_{\diamond t}\}_{t=1}^T$, we define the context-aware energy $E_c(S)$ by

$$E_c(S) = -\sum_{t=1}^T S^T W_{\diamond t} S_{\diamond t}, \qquad (4)$$

$W_{\diamond t}$ captures the interaction strength between dictionary atoms in orientation t. For instance, to the current patch x_i, $W_{\diamond t}(m, n) = 0$ indicates $S_i(m)$ and $S_{i \diamond t}(n)$ tend to be independent; $W_{\diamond t}(m, n) > 0$ indicates $S_i(m)$ and $S_{i \diamond t}(n)$ tend to

be activated simultaneously; $W_{\diamond t}(m, n) < 0$ indicates $S_i(m)$ and $S_{i\diamond t}(n)$ tend to be mutually exclusive. We will introduce how to compute $W_{\diamond t}$ later in this section.

Meanwhile, the sparsity penalty energy $E_s(S)$ is taken into account:

$$E_s(S) = -S^T b, \tag{5}$$

where $b = [b_1, b_2, ..., b_m]^T$ is a vector of model parameters, and b_i is associated with the dictionary atom, $b_i < 0$ favors $S_i = -1$. The total energy for each sparsity pattern is the sum of the context-aware energy and the sparsity energy, *i.e.*, $E_{total} = E_c(S) + E_s(S)$. The prior probability can then be formalized using the total energy,

$$\Pr(S) \propto \exp\left(-E_{total}\right)$$
$$\propto \exp\left(S^T \left(\sum_{t=1}^{T} W_{\diamond t} S_{\diamond t} + b\right)\right). \tag{6}$$

Let $\tilde{W} = [W_{\diamond 1}, \ldots, W_{\diamond J}]$, and $\tilde{S} = \left[(S_{\diamond 1})^T, \ldots, (S_{\diamond J})^T\right]^T$, then eq.(6) can be expressed in a clearer form,

$$\Pr(S) = \frac{1}{Z(\tilde{W}, b)} \exp\left(S^T \left(\tilde{W}\tilde{S} + b\right)\right), \tag{7}$$

where $\tilde{W}, b$ are model parameters, and $Z(\tilde{W}, b)$ is the function for normalization. It shows compared with conventional sparsity priors, the proposed prior model places more emphasis on the dependencies of atoms in the spatial context.

For the above new prior, the model parameters including $\tilde{W}, b$, and $\{\sigma_{\gamma,i}^2\}_{i=1}^m$ should be estimated. $\sigma_{\gamma,i}^2$ stands for variance of each nonzero coefficient γ_i. Given $\mathcal{X} = \{x^k, S^k, \gamma^k, \tilde{S}^k\}_{k=1}^K$ as examples sampled from the model, we suggest using the Maximum Likelihood Estimation (MLE) for learning the model parameters $\theta = [\tilde{W}, b, \{\sigma_{\gamma,i}^2\}_{i=1}^m] \in \Theta$. Mathematically, we have

$$\hat{\theta}_{ML} = \arg\max_\theta \Pr\left(\mathcal{X}|\theta\right) = \arg\max_\theta \sum_{i=1}^m \mathcal{L}(\sigma_{\gamma,i}^2) + \mathcal{L}(\tilde{W}, b), \tag{8}$$

where

$$\mathcal{L}(\sigma_{\gamma,i}^2) = \frac{1}{2}\sum_{k=1}^K f_i^k,$$

$$\mathcal{L}(\tilde{W}, b) = \frac{1}{2}\sum_{k=1}^K \left(S^k\right)^T \left(\tilde{W}\tilde{S}^k + b\right) - K \ln Z(\tilde{W}, b), \tag{9}$$

are log-likelihood functions for the model parameters and

$$f_i^k = \begin{cases} \dfrac{\left(\gamma_i^k\right)^2}{\sigma_{\gamma,i}^2} + \ln(\sigma_{\gamma,i}^2), & S_i^k = 1, \\ 0, & S_i^k = -1. \end{cases} \tag{10}$$

For the estimation of variances, a closed-form estimator is obtained by:

$$\hat{\sigma}_{\gamma,i}^2 = \frac{\sum_{k=1}^K (\gamma_i^k)^2 \cdot q_i^k}{\sum_{k=1}^K q_i^k}, \tag{11}$$

where

$$q_i^k = \begin{cases} 1, & S_i^k = 1, \\ 0, & S_i^k = -1. \end{cases}$$

However, ML estimation of $\tilde{W}$ and b is computationally intensive due to the exponential complexity in m associated with the partition function $Z(\tilde{W}, b)$. We adopted an efficient algorithm [9] using the MPL estimation and sequential sub-space optimization (SESOP) method to tackle the problem.

2.4. Image Reconstruction

The prior model proposed in the previous subsection is defined in a patch-wise scheme. It is enforced over the local neighborhood range of each patch. In fact, the neighboring sparsity patterns $\tilde{S}$ are always unknown when addressing the sparsity pattern recovery for one single patch. Meanwhile, when dealing with an arbitrary size image, it is necessary to extend the local prior to a global one as in [10, 11] and we incorporate the context-aware sparsity prior into the MRFs framework.

For an input degraded image X of arbitrary size, we first break it into overlapped small patches $\{x^k\}_{k=1}^K$. Each patch x^k has a corresponding high-quality patch y^k, and the "true" sparsity pattern of y^k is denoted as S^k. $\mathcal{S} = \{S^k\}_{k=1}^K$ represents the whole set of sparsity patterns. We introduce an 8-connected MRFs to model the relationships among the degraded patches and their corresponding high-quality patches. Based on the MRFs model, we define three types of potential functions corresponding to the likelihood term $\phi(S^k, x^k)$, sparsity term $\eta(S^k)$ and context-aware term $\psi(S^k, S^p)$,

$$\phi(S^k, x^k) \propto \Pr(x^k|S^k),$$
$$\eta(S^k) \propto \exp\left((S^k)^T b\right), \tag{12}$$
$$\psi(S^k, S^p) \propto \exp\left((S^k)^T W_{\diamond t} S_{\diamond t}^k\right),$$

which use the fact that patch y^p is adjacent to y^k in the t-th orientation. Once the potential functions are determined, the MRFs with homogeneous potentials could be written as

$$\Pr(\mathcal{S}, X) \propto \prod_k \eta(S^k)\phi(S^k, x^k) \prod_{k,p} \psi(S^k, S^p). \tag{13}$$

$\phi(S^k, x^k)$ corresponds to the likelihood probability $\Pr(x^k)$. Therefore, the complete set of sparsity pattern $\mathcal{S}$ in the MRFs can be optimally estimated by maximizing the joint probability of MRFs,

$$\max \Pr(\mathcal{S}, X) = \max \sum_{k=1}^K \left(\ln \Pr\left(x^k|S^k\right) + \ln \Pr\left(S^k\right)\right). \tag{14}$$

Since the parameters are calculated, one way to compute the global optimal configuration for the MRFs model in (14) is to provide a set of possible candidates for each node, then

approximately solve it by the Belief Propagation algorithm. However, since the number of possible configurations of each node is exponential to the number of the dictionary atoms (i.e., there are 2^m possible candidates for S), it is computationally intractable in practice. Thus we present an approximated numerical solution that iteratively recovers the sparsity pattern of each patch, as in the Gauss-Seidel iterative method.

In the proposed algorithm, all the patches are processed in raster-scan order in an image, i.e., from left to right and top to bottom. When processing the current center patch, all sparsity patterns of the neighboring patches $\tilde{S}$ are utilizing the latest updated value and kept fixed during the recovery of sparsity pattern for center patch. Due to the overlapping of extracted patches, the updated sparsity pattern of the current patch is immediately used in the processing of next neighboring patch. The procedure is performed repeatedly to propagate the contextual information among all the nodes.

The above simplification for solving the whole set of sparsity patterns of the MRFs can be viewed as a block-coordinate method, in which when updating one single sparsity pattern, the others are known and fixed. We adopt a greedy algorithm as an approximate MAP estimation for computing sparsity patterns. The greedy algorithm starts with an initialization with $S_i = -1, \forall i$, and then iteratively changes the value of entry S_i to 1 that makes the posterior probability of S with the biggest growth comparing to all other candidates. The iteration stops until the the posterior probability reaches a local optimal value.

Algorithm 1: MRF-based Image Recovery Algorithm

Input: Noisy observations $\{x^k\}_{k=1}^K$, dictionary $\tilde{D}$, noise variance σ^2, model parameters $\theta = \left[\tilde{W}, b, \upsilon = \{\sigma_{\gamma,i}^2\}_{i=1}^m \right]$, initialization $\mathcal{S}^{(0)}$, maxPass.
Output: Recovery of HR image y.
$p = 0$;
while $p < maxPass$ **do**
    $p = p + 1$;
    for *every patch x^k in raster-scan order* **do**
        Collect the sparsity patterns of neighboring patches, $\tilde{S}^k$;
    end
end
return $\hat{S} = \{\hat{S}^k\}_{k=1}^K$.
$\hat{\gamma}_S = \arg \max_{\gamma_s} \Pr(\gamma_s | x, \hat{S}) = \mathbf{Q}_S^{-1} \tilde{D}_S^T x,$
$y = \tilde{D}_S \hat{\gamma}_S = \tilde{D}_S \mathbf{Q}_S^{-1} \tilde{D}_S^T x.$

With sparsity patterns known, we can estimate the sparse codes and reconstruct the HR image as follows:

$$\hat{\gamma}_S = \arg \max_{\gamma_s} \Pr(\gamma_s | y, \hat{S}) = \mathbf{Q}_S^{-1} \tilde{D}_S^T x,$$
$$y = \tilde{D}_S \hat{\gamma}_S = \tilde{D}_S \mathbf{Q}_S^{-1} \tilde{D}_S^T x. \quad (15)$$

where the nonzero coefficients in γ are denoted as γ_S, and the corresponding atoms in $\tilde{D}$ which participate in the representation γ_S are grouped into a sub-dictionary denoted by $\tilde{D}_S$. Σ_S is a $k \times k$ diagonal matrix in which the diagonal elements are the corresponding variances $\sigma_{\gamma,i}^2$ of the nonzero coefficients γ_i, and k is the total number of the nonzero coefficients in γ. $\mathbf{Q}_S = \tilde{D}_S^T \tilde{D}_S + \sigma^2 \Sigma_S^{-1}$.

The pseudocode of the MRF-based image recovery algorithm is summarized in Algorithm 1.

3. EXPERIMENTAL RESULTS

To evaluate the efficiency of the proposed method, we conduct experiments of $3\times$ super resolution on several test sets. The LR input images are generated from the original HR images by downsampling with bicubic method by the scaling factor, and contaminated by additive Gaussian noise with standard deviation $\sigma_n = 1$.
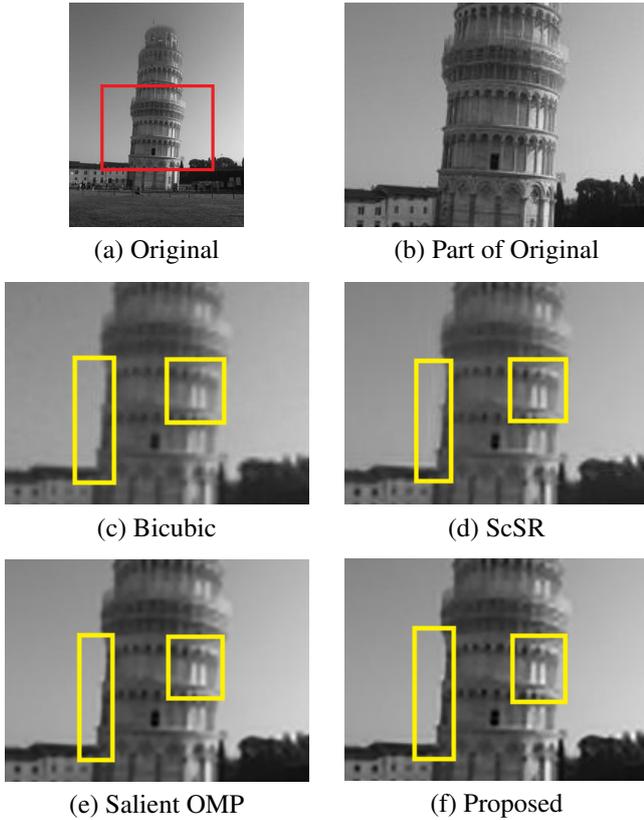

Fig. 4. The local neighborhood system of patch x_i with a spatial configuration of eight different orientations.

Table 1. PSNR (dB) Comparison of Different Methods of $3\times$ SR on Test Images.

Images	Bicubic	ScSR	Salient OMP	Proposed
Liberty	23.03	23.50	23.57	**23.60**
Relic	24.12	24.58	**24.69**	24.68
Palace	23.34	23.81	23.89	**23.92**
Horse	27.61	27.92	28.11	**28.13**
Colosseum	22.21	22.70	22.79	**22.81**
Tower	29.52	29.95	30.08	**30.09**
Average	24.97	25.41	25.52	**25.54**

We test the proposed method on six image sets, *Liberty*, *Relic*, *Palace*, *Horse*, *Colosseum*, and *Tower*, all collected from the internet(released on our website[2]). Half of them as training examples, the rest are used as test images (see Fig.4). For each database, a general dictionary and a salient dictionary are learned separately. The LR patch size is 3×3 and therefore HR patch size is 9×9, and the overlaps between patches are [2, 2] and [6, 6] for LR and HR patches,

---

[2]http://www.icst.pku.edu.cn/course/icb/SalientSR.html

(a) Original   (b) Part of Original

(c) Bicubic   (d) ScSR

(e) Salient OMP   (f) Proposed

Fig. 5. Comparison of different methods on image *Tower*.

respectively. We compare our method with the baseline bicubic method, ScSR [3], and OMP with salient dictionary (see Table.1). ScSR is one of the state-of-the-art SR algorithms, and the proposed method shows 0.1 0.2dB PSNR gain over it. Then, for the integrity of the whole verification process, we incorporate saliency into the traditional OMP-based SR method to develop the saliency OMP to demonstrate the effectiveness of saliency segmentation.

At the same time, we show subjective results on test set *Tower*. Fig.4 shows zoomed comparison of the highlighted part in the original image by different methods. Compared with ScSR, Fig.4(e) and (f) significantly reduce artifacts along the tower edges thanks to the salient dictionary. Meanwhile, owing to the context-aware sparse decomposition, more structural information is recovered (see details of the tower columns, best view on screen).

4. CONCLUSION

In this work, based on sparse representation SR framework, we focus on how to make the most of the underlying structural information in images. Considering the property of salient regions in images, we propose a saliency based dictionary learning pattern. Another contribution of this work is we incorporate context-aware sparse decomposition to model dependen-

cies between dictionary atoms of adjacent patches. Experimental results show the proposed method outperforms other methods in both objective and subjective quality.

5. REFERENCES

[1] K. Engan, S. Ease, and J. Husoy, "Multi-frame compression: Theory and design," *Signal Processing*, vol. 80, no. 10, pp. 2121–2140, Oct. 2000.

[2] M. Aharon, M. Elad, and A. Bruckstein, "K-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[3] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[4] M. Protter, I. Yavneh, and M. Elad, "Closed-form mmse estimation for signal denoising under sparse representation modeling over a unitary dictionary," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3471–3484, July 2010.

[5] N. G. Sadaka and L. J. Karam, "Efficient super-resolution driven by saliency selectivity," in *IEEE International Conference on Image Processing (ICIP)*, Sept. 2011, pp. 1197–1200.

[6] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[7] S.S. Chen, D.L. Donoho, and M.A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, Jan. 2001.

[8] R. Achanta, F. Estrada, P. Wils, and S. Susstrunk, "Salient region detection and segmentation," *International Conference on Computer Vision Systems (ICVS)*, pp. 66–75, Springer, 2008.

[9] A. Hyvarinen, "Consistency of pseudolikelihood estimation of fully visible boltzmann machines," *Neural Computation*, vol. 18, no. 10, pp. 2283–2292, 2006.

[10] S. Roth and M.J. Black, "Fields of experts: A framework for learning image priors," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2005, vol. 2, pp. 860–867.

[11] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.