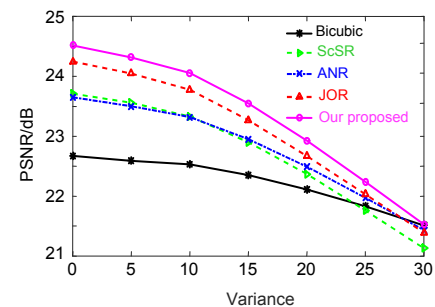# MODIS image super-resolution via learned topic dictionaries and regression matrices

Deng Zuo[1], Randi Fu[1]*, Wei Jin[1] and Caifen He[2]

[1]Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China; [2]Zhenhai District Meteorological Bureau, Ningbo 315202, China

Abstract: Moderate resolution imaging spectroradiometer (MODIS) imaging has various applications in the field of ground monitoring, cloud classification and meteorological research. However, the limitations of the sensors and external disturbance make the resolution of image still limited in a certain level. The goal of this paper is to use a single image super-resolution (SISR) method to predict a high-resolution (HR) MODIS image from a single low-resolution (LR) input. Recently, although the method based on sparse representation has tackled the ill-posed problem effectively, two fatal issues have been ignored. First, many methods ignore the relationships among patches, resulting in some unfaithful output. Second, the high computational complexity of sparse coding using l_1 norm is needed in reconstruction stage. In this work, we discover the semantic relationships among LR patches and the corresponding HR patches and group the documents with similar semantic into topics by probabilistic Latent Semantic Analysis (pLSA). Then, we can learn dual dictionaries for each topic in the low-resolution (LR) patch space and high-resolution (HR) patch space and also pre-compute corresponding regression matrices for dictionary pairs. Finally, for the test image, we infer locally which topic it corresponds to and adaptive to select the regression matrix to reconstruct HR image by semantic relationships. Our method discovered the relationships among patches and pre-computed the regression matrices for topics. Therefore, our method can greatly reduce the artifacts and get some speed-up in the reconstruction phase. Experiment manifests that our method performs MODIS image super-resolution effectively, results in higher PSNR, reconstructs faster, and gets better visual quality than some current state-of-art methods.

Keywords: MODIS; super-resolution; sparse representation; sparse coding; regression matrix

DOI: 10.3969/j.issn.1003-501X.2017.10.003   **Citation:** *Opto-Elec Eng*, 2017, **44**(10): 957−965

1 Introduction

There are many factors which cause the resolution lower in the imaging process of remote sensing image, such as atmospheric disturbance, frequency aliasing[1]. However, spatial resolution is an important property of remote image sensing [2]. As an important branch of remote sensing, the moderate resolution imaging spectroradiometer (MODIS), mounted on Terra and Aqua satellites, is an important instrument for observing global biological and physical processes in the Earth Observation System (EOS) program. There are 36 medium resolution levels (0.25 μm~1 μm) in the MODIS spectral band[3]. It is widely used in the fields of ground detection, cloud classification

and climate research because it contains rich information. However, due to sensor limitations and external interference, MODIS image resolution is limited to a certain level. Therefore, using super-resolution technology to improve resolution of the MODIS image has a great practical significance and application meaning.

The conventional SISR approaches can be mainly divided into two catalogs: regularized methods[4] and interpolation methods[5]. However, the performance of regularization methods degrades rapidly with the increase of magnification factor[6], and the interpolation methods do not take prior knowledge into account so that the enlarged image cannot increase any information in nature. Recently, another class of super-resolution schemes, examples-based methods[7-9], which use a learned co-occurrence prior between LR and HR image patches space to predict a HR image. Chang et al. proposed a neighbor embedding[10] (NE) super-resolution algorithm

which adopted the philosophy of local linear embedding (LLE)[11] from manifold learning, assuming similar local geometry between the two manifolds in the HR image patch space and LR image patch space.

The NE algorithm maps the local geometry of LR image patch space to the HR image patch space and predicts HR image patch as a linear combination of neighbors. However, due to over-or under-fitting, this strategy often results in smooth effects. In very recent years, the sparse representation based SISR has attracted increasing interests and several works are proposed in Refs.[7-9]. In Refs.[7-8], Yang et al. adopt the compressive sensing[12](CS) theory and extend the NE to the sparse representation scheme, which assumes each image patch from test image considered can be well represented using a linear combination of a few atoms from a sparse dictionary. Considering that a large number of example patches are needed to learn HR and LR dictionaries for accurate reconstruction, Zeyde et al.[13] improved the Yang's framework to reduce the computational complexity of dictionary learning by K-SVD algorithm[14]. In Ref.[15], Timofte et al. proposed an ANR super-resolution algorithm, learned single dictionary pair and regression matrices anchored to the dictionary atoms obtained better quality and largely speed up in the reconstruction phase by learning the regression matrices in the full HR and LR spaces. Nevertheless, the relationships among the patches with repetitive structures are still ignored, which will result in noticeable reconstruction artifacts.

With the above consideration, in this paper we propose to discover semantic relationships among patches in the documents by pLSA[16] and group the documents with similar semantic prior into topics, and then, apply the learned prior to LR image for high performance reconstruction. People are sensitive to the high frequency detail part of image. In Ref.[17], the authors have demonstrated the extracted high frequency detail images are sparser than the original images and compressive sensing shows that the sparser signal will have higher reconstruction accuracy. Therefore, we select bilateral filter as a non-linear filter to extract the detail part of MODIS images, which viewed as the training images. The flowchart of proposed method is illustrated in Fig. 1. In the learning phase, we divided the detail part into documents and documents into patches and grouped documents into different topics by their semantic relationships. For each
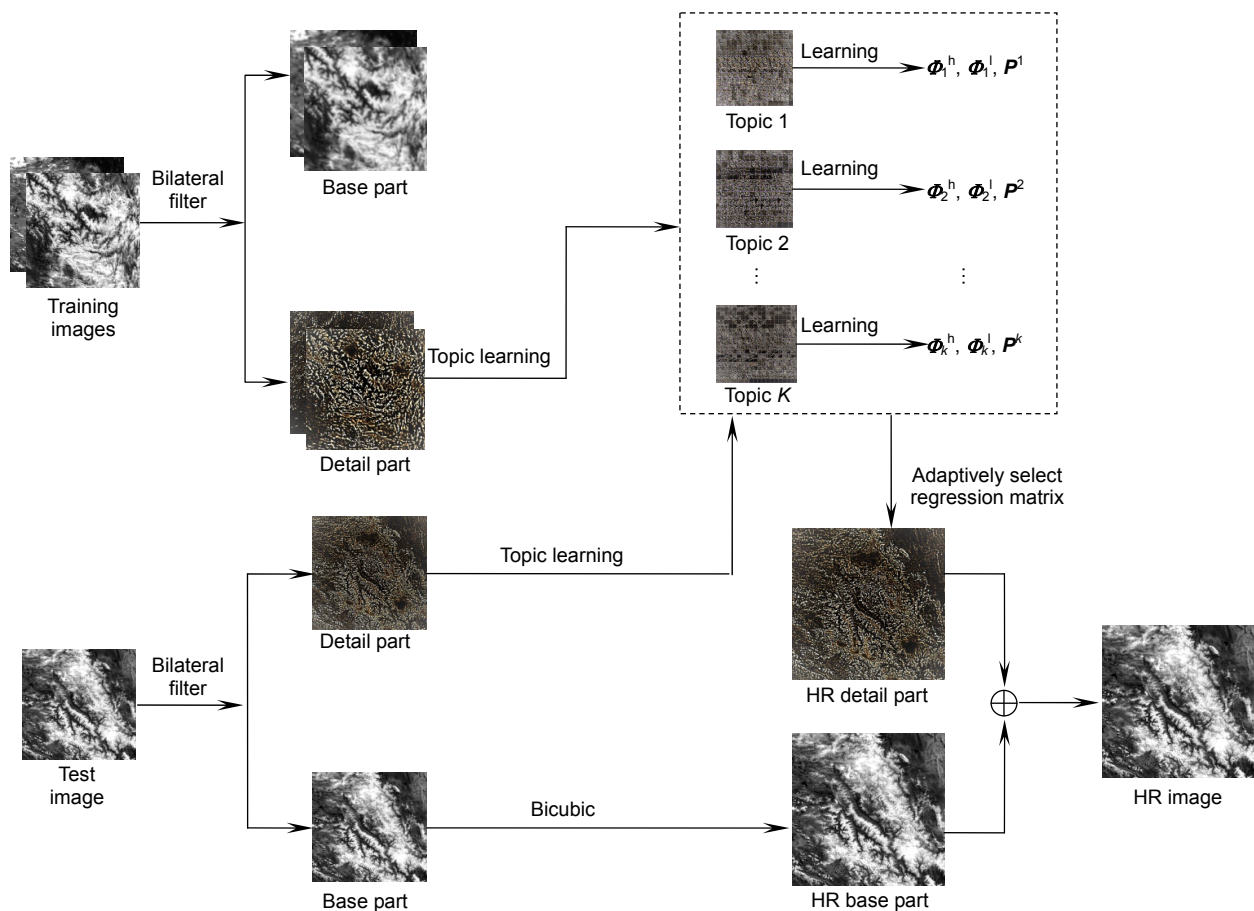

Fig. 1 Block diagram of the proposed algorithm.

topic, we adopt the method of Zeyde et al.[13] to learn a pair of dictionaries and calculate the regression matrix for the dictionaries. In the reconstruction phase, we also divided the LR image into base part and detail part by bilateral filter. In this part, we reconstruct HR detail part image and HR base part image respectively. For the LR detail part, we use the pLSA[16] to discover locally which topic it corresponds to and adaptively select the regression matrix for reconstruction. In section 2.3, we explain how to calculate the regression matrices and why the regression matrices can get some speed-up in the reconstruction phase. At the other part, we simply used bi-cubic interpolation to restore it. Finally, an HR MODIS image can be restored by combining HR detail part and HR base part. The experiment shows that the proposed algorithm can not only reconstruct accurately but also get some speed-up in the reconstruct stage over some existing state-of-art methods.

2 Proposed framework

Both Yang's and Zeyde's single image super-resolution framework learn the one-to-one correspondence between LR and HR patch pairs simply and just predict a HR patch accordingly from a LR patch. But infinitely many HR patches will result in the same LR patch when blurred and down-sampled. This is an extremely ill-pose problem. Therefore, we group the LR patches with similar semantic and the corresponding HR patches into topics in the training stage and find the HR patch with the most similar semantic from all possible HR patches for a given LR patches in the reconstruct stage by pLSA[16]. With the above processing, we can get the optimal reconstruction for the HR detail part. In the subsequent section, we will elaborate our work.

2.1 Extract detail part by bilateral filter

The detail part contains the rich high frequency information of image. We select the bilateral filter as a non-linear filter to extract the base part from MODIS image. Then, the detail part can be got by the difference between original MODIS image and base part. Bilateral filter contains two functions. One is determined by geometric spatial distance and the other is determined by the difference of pixel values. The output of the bilateral filter is:

$$X_b(i,j) = \frac{1}{\gamma(i,j)} \sum_{(i',j') \in S_{i,j}} g_s(i-i', j-j') *$$
$$g_r(X_{in}(i,j) - X_{in}(i',j'))X_{in}(i',j'), (1)$$

where $(i',j') \in S_{i,j}$ means pixel (i',j'), (i,j) are adjacent pixels in the image, and $S_{i,j}$ represents the template domain. The $\gamma(i,j)$ is the normalized parameter, which can be written as

$$\gamma(i,j) = \sum_{(i',j') \in S_{i,j}} g_s(i-i', j-j') *$$
$$g_r(X_{in}(i,j) - X_{in}(i',j')), \qquad (2)$$

where g_s is a standardized Gaussian kernel function,

g_r is a Gaussian function in the intensity domain, $X_{in}(i,j)$ is the original image. Detail part $X_d(i,j)$ can be got by following equation:

$$X_d(i,j) = X_{in}(i,j) - X_b(i,j). \qquad (3)$$

2.2 Topic learning model

The topic learning model can discover the semantic relationships of different possible HR patches corresponding to a given LR patch. Therefore, our intention is to discover the topics locally for a given image set. In order to apply the topic learning model to MODIS image super-resolution, we divide the training images (detail parts) into N local regions $D = (d_1, d_2, \cdots, d_N)$ and define each local region as a document. N documents are divided into M words, which represent patches. Topic $z \in Z = \{z_1, z_2, \cdots, z_k, \cdots, z_K\}$ is even higher level concept that group different documents according to the co-occurrences of different words within and across the documents[18]. In our proposed method, we group the patches in documents into topics by topic learning model and learn a pair of high resolution and low resolution dictionaries for each topic.

For a given document, we can determine which topic it belongs to by the pLSA[16]. The collection of documents and collection of words can be represented in a $M \times N$ co-occurrence table $f(w_i, d_j)$, which stores the frequency of occurrences of the word w_i in the document d_j, and N and M are the number of documents and words size, respectively. Suppose there are K latent topics z_k associated with the collection of documents and there are certain topic proportions corresponding to each document. PLSA is a topic model, which can discover the latent topic z_k with a certain probability for the given document d_j and select the corresponding words w_i from the topic z_k with a certain probability. The pLSA topic model can be represented as follows:

$$P(w_i, d_j) = P(d_j)P(w_i/d_j), \qquad (4)$$

$$P(w_i/d_j) = \sum_{k=1}^{K} P(z_k/d_j)P(w_i/z_k), \qquad (5)$$

where $P(w_i, d_j)$ is the joint probability of words and documents, $P(d_j)$ is the prior probability of the documents which assumed to follow an uniform distribution, $P(z_k/d_j)$ is the probability of latent topic z_k occurring in document d_j, $P(w_i/d_j)$ is the conditional probability and $P(w_i/z_k)$ is the probability of word w_i occurring in a particular topic z_k.

Fitting the model involves determining the topic vectors $P(w_i/z_k)$ which are common to all documents and the document specific mixture coefficients $P(z_k/d_j)$. According to the maximum likelihood estimation principle, the parameters of the pLSA model can be obtained by calculating the maximum likelihood function:

$$L = \sum_{i=1}^{M} \sum_{j=1}^{N} f(w_i, d_j) \log P(w_i, d_j). \qquad (6)$$

A classical algorithm to solve maximum likelihood estimation function is EM algorithm. After initializing with

a random value, the E-step and M-step are used to iterative calculation alternately until convergence. In the E-step, the posterior probability of each latent variable is calculated:

$$P(z_k/w_i, d_j) = \frac{P(z_k)P(d_j/z_k)P(w_i/z_k)}{\sum_{l=1}^{K} P(z_l)P(d_j/z_l)P(w_i/z_l)} . \quad (7)$$

In the M-step, the parameters are estimated:

$$P(w_i/z_k) = \frac{\sum_{j=1}^{N} f(w_i, d_j)P(z_k/w_i, d_j)}{\sum_{t=1}^{M}\sum_{j=1}^{N} f(w_t, d_j)P(z_k/w_t, d_j)} , \quad (8)$$

$$P(d_j/z_k) = \frac{\sum_{i=1}^{M} f(w_i, d_j)P(z_k/w_i, d_j)}{\sum_{j=1}^{N}\sum_{t=1}^{M} f(w_t, d_j)P(z_k/w_t, d_j)} , \quad (9)$$

$$P(z_k/d_j) = \frac{P(d_j/z_k)P(z_k)}{\sum_{l=1}^{K} P(d_j/z_l)P(z_l)} . \quad (10)$$

After each iteration of equation in the M-step, we can learn the $P(z_k/d_j)$ which represents the mixture coefficients of each document. The maximum value $P(z_k/d_j)$ of the document can be used to determine the document topic assignment. In Fig. 2, we displayed the documents under three different typical topics trained by the detail part of MODIS images. It is clear that the pLSA[16] method clusters the documents with similar semantic together under topics. Therefore, dividing the documents into different topics and learning dual dictionaries for each topic will effectively represent structures of image. $P(w_i/z_k)$ and $P(z_k)$ are stored for future EM algorithm that would be used for the test documents topic prediction.

2.3 Topic dictionary learning and regression matrix learning

For each topic z_k, there are HR detail part patches set X_h^k and the corresponding LR detail part patches set Y_1^k, we learn a dual dictionaries $\{\Phi_h^k, \Phi_1^k\}$ over the words(patches) under topic. Then the LR dictionary Φ_1^k can be learned as follow:

$$\{\Phi_1^k, A^k\} = \arg\min_{\Phi_1^k, A^k} \left\| Y_1^k - \Phi_1^k A^k \right\|_2^2 =$$

$$\arg\min_{\Phi_1^k, A^k} \sum_s \left\| y_{1,s}^k - \Phi_1^k \alpha_s \right\|_2^2 , \quad (11)$$

$$\text{s.t.} \left\| \alpha_s \right\|_0 \leq Q,$$

where Q is the sparsity, $y_{1,s}^k$ is a certain patch in Y_1^k and α_s is the sparse representation coefficient. We use K-SVD algorithm[14] and OMP algorithm[19] to solve the Eq. (11).

Sparse representation theory supposed that HR patches X_h^k and the corresponding LR patches Y_1^k have same sparse representation coefficient. Therefore, the HR dictionary Φ_h^k can be learned as follow equation:

$$\Phi_h^k = \arg\min_{\Phi_h^k} \sum_s \left\| x_{h,s}^k - \Phi_h^k \alpha_s \right\|_2^2 =$$

$$\arg\min_{\Phi_h^k} \left\| X_h^k - \Phi_h^k A^k \right\|_2^2 , \quad (12)$$

where $x_{h,s}^k$ represents a certain patch in X_h^k. The solution of Eq. (12) is given by the following Pseudo-Inverse expression:

$$\Phi_h^k = X_h^k (A^k)^+ = X_h^k A^k [A^k (A^k)^{\mathrm{T}}]^{-1} . \quad (13)$$

Timofte et al. proposed ANR[15] super-resolution algorithm, and learned regression matrices anchored to the LR-HR dictionary atoms can largely reduce the computational complexity in the reconstruction stage. Therefore, using a similar framework to Timofte et al, once the LR-HR topic dictionary pair subsets $\{\Phi_h^k, \Phi_1^k\}_{k=1}^K$ are estimated, multiple regression matrices are computed by Collaborative Representation[20] to predict HR patch space from LR patch space.

Unlike most sparse coding super-resolution methods that solve the sparse coefficients by ℓ_0 norm or ℓ_1 norm regularization term, we select Collaborative Representation[20] with ℓ_2 norm regularized least squares regression to compute the regression matrices for each LR-HR topic dictionary pair subsets $\{\Phi_h^k, \Phi_1^k\}_{k=1}^K$. The closed-form solution problem becomes:
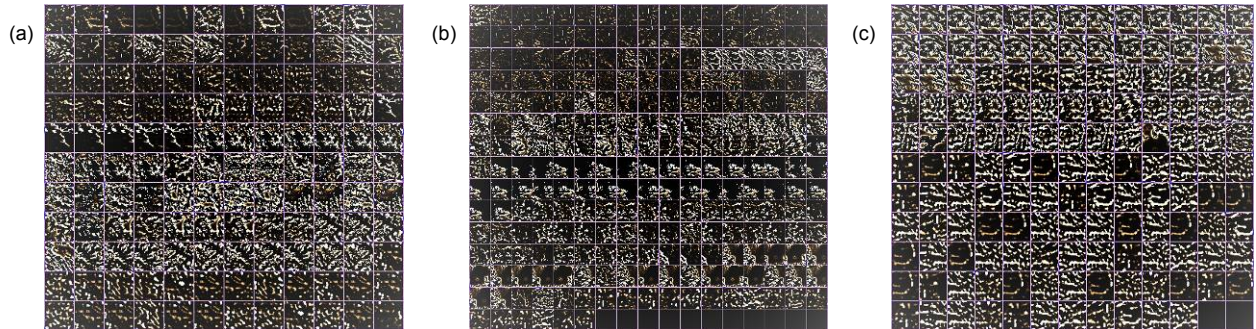

Fig. 2 Collection of documents under different topics. (a) Topic 8. (b) Topic 9. (c) Topic 10.

$$\boldsymbol{\beta} = \min_{\boldsymbol{\beta}} \left\| \boldsymbol{y} - \boldsymbol{\Phi}_1^k \boldsymbol{\beta} \right\|_2^2 + \lambda \left\| \boldsymbol{\beta} \right\|_2 , \qquad (14)$$

where $\boldsymbol{y}$ is a certain LR patch, $\boldsymbol{\Phi}_1^k$ is the kth LR dictionary subset whose semantic relationship is most similar to $\boldsymbol{y}$, $\boldsymbol{\beta}$ is the coefficient vector of $\boldsymbol{y}$ over $\boldsymbol{\Phi}_1^k$, and $\lambda > 0$ is a weighting parameter which allows us to alleviate the singularity problems and stabilizes the solution. The closed-form solution is given by:

$$\boldsymbol{\beta} = ((\boldsymbol{\Phi}_1^k)^{\mathrm{T}} \boldsymbol{\Phi}_1^k + \lambda \boldsymbol{I})^{-1} (\boldsymbol{\Phi}_1^k)^{\mathrm{T}} \boldsymbol{y} . \qquad (15)$$

Assuming that LR patch $\boldsymbol{y}$ and the corresponding HR patch $\boldsymbol{x}$ share the same coefficient vector $\boldsymbol{\beta}$ over $\boldsymbol{\Phi}_1^k$ and $\boldsymbol{\Phi}_h^k$ respectively, the desired output HR patch can be obtained by:

$$\boldsymbol{x} = \boldsymbol{\Phi}_h^k \boldsymbol{\beta} = \boldsymbol{\Phi}_h^k ((\boldsymbol{\Phi}_1^k)^{\mathrm{T}} \boldsymbol{\Phi}_1^k + \lambda \boldsymbol{I})^{-1} (\boldsymbol{\Phi}_1^k)^{\mathrm{T}} \boldsymbol{y} . \qquad (16)$$

By observing the equation (16), we can obtain the regression matrix:

$$\boldsymbol{P}^k = \boldsymbol{\Phi}_h^k ((\boldsymbol{\Phi}_1^k)^{\mathrm{T}} \boldsymbol{\Phi}_1^k + \lambda \boldsymbol{I})^{-1} (\boldsymbol{\Phi}_1^k)^{\mathrm{T}} . \qquad (17)$$

By comparing Eq. (16) with Eq. (17), the corresponding regression matrices can be calculated once we get the topic dictionaries. The regression matrices can be computed offline. As we know, the methods based on sparse representation will calculate sparse coefficient for each test patch respectively. It is time consuming. Our method only needs discovering the semantic relationship by pLSA[16] and adaptively selecting the regression to predict HR patch for an input LR patch. It means that we can avoid calculating the sparse coefficient. Therefore, our method can get some speed-up in the reconstruction phase.

2.4 Summary of proposed algorithm

The newly proposed SR algorithm consists of two phases: learning phase and reconstruction phase. Learning phase: we applied the pLSA topic learning model to discover the latent topics from an image database and learned the dual dictionaries and regression matrix for each topic. Reconstruction phase: desired HR image is predicted from input LR image by topic assignment for each patch and adaptively selecting the regression matrix to obtain the HR patch. The learning phase and the reconstruction phase of our algorithm are summarized in Algorithm 1 and Algorithm 2, respectively.

Algorithm 1: topic learning, dictionary learning and regression matrix learning

1) Input: LR and HR MODIS images, scaling factor S.

2) Preprocessing: Using bilateral filter $X_{\mathrm{b}}(i, j)$ to extract detail parts from the LR and HR MODIS images.

3) Dividing the LR and HR detail parts into overlapping documents with size $m \times m$ and the documents into overlapping patches with size $n \times n$.

4) For $o = 1$; $o = M$; $o{+}{+}$ do.

5) Appling the pLSA topic learning model to the documents and inferring the topic $\boldsymbol{z}_k$ for each document $\{\boldsymbol{d}_j\}_{j=1}^N$ by EM algorithm.

6) The E-step and M-step of EM algorithm are used to iterate the Eqs. (7)~(10) until convergence.

7) Obtain the maximum mixture coefficients $P(\boldsymbol{z}_k / \boldsymbol{d}_j)$ of document $\boldsymbol{d}_j$.

8) Group the documents into topics by maximum mixture coefficients $P(\boldsymbol{z}_k / \boldsymbol{d}_j)$.

9) End for.

10) Using K-SVD to learn dictionaries $\{\boldsymbol{\Phi}_h^k, \boldsymbol{\Phi}_1^k\}_{k=1}^K$ for each topic by Eqs. (11) and (13) and learning regression matrices $\{\boldsymbol{P}^k\}_{k=1}^K$ by Eq. (17).

11) Storing document-topic distribution and the regression matrices $\{\boldsymbol{P}^k\}_{k=1}^K$ that would be used for the test image super-resolution.

Algorithm 2: the reconstruction phase of the proposed algorithm

1) Input: the test LR MODIS image $\boldsymbol{Y}$, scaling factor S, document-topic distribution regression matrices $\{\boldsymbol{P}^k\}_{k=1}^K$.

2) The test LR MODIS image $\boldsymbol{Y}$ is decomposed into LR detail part $\boldsymbol{Y}_{\mathrm{d}}^1$ and LR base part $\boldsymbol{Y}_{\mathrm{b}}^1$ by the bilateral filter $X_{\mathrm{b}}(i, j)$.

3) Dividing the LR detail part $\boldsymbol{Y}_{\mathrm{d}}^1$ into overlapping documents with size $m \times m$ and the documents into overlapping patches with size $n \times n$.

4) Appling the pLSA topic learning model to the documents and inferring the topic $\boldsymbol{z}_k$ for each document $\boldsymbol{d}$ by EM algorithm.

5) Obtained the maximum mixture coefficients $P(\boldsymbol{z}_k / \boldsymbol{d})$ which is determined topic assignment of document $\boldsymbol{d}$.

6) Adaptively selecting the corresponding regression matrix $\boldsymbol{P}^k$ for the inferred topic of a document to predict the HR patches in the document. Then make an average of all the HR patches in overlapping region of the overlapping documents to obtain the HR detail part $X_{\mathrm{d}}^{\mathrm{h}}$.

7) For HR base part $X_{\mathrm{b}}^{\mathrm{h}}$, it can be restored simply by bi-cubic interpolation.

8) Combining the HR detail part $X_{\mathrm{d}}^{\mathrm{h}}$ with the HR base part $X_{\mathrm{b}}^{\mathrm{h}}$ to reconstruct the HR MODIS image $X = X_{\mathrm{d}}^{\mathrm{h}} + X_{\mathrm{b}}^{\mathrm{h}}$.

We collect a number of sharp MODIS image as the HR images. We generate corresponding LR images as follows:

1) For each image in the collection, using a Gaussian blur operator with small kernel (size 3×3 and $\sigma = 0.8$) to obtain blur versions.

2) Down-sample each blurred image with resolution factor S.

3) Then scaled up back to the original size by bi-cubic interpolation. Thus original HR and corresponding LR images are same size.

For the given HR and LR images, we use the bilateral filter to extract the detail parts as the training images database and then divide the HR detail parts and corresponding LR detail parts into overlapping documents and the documents into overlapping patches.

3　Experimental results

In this part, in order to verify the performance of our proposed method, we have implemented the experiment by Matlab R2010b on windows 10 with 16 GB RAM and a 2.6-GHz intel processor. The training samples were obtained from the MODIS images generated in the channels 1~4, 6, 7, 17~19, 22~26, 31~36 of the MODIS data. The test images were obtained from 7:57 pm on March 2, 2016. AQUA satellite reception MODIS 1B data, the area is 97.8151 degrees east longitude to 134.6207 degrees east longitude, 46.3421 degrees north latitude to 97.8151 degrees. In our experiments, we have divided the training images into 20000 documents with size 50×50 and the patches size 7×7. In order to reduce the computational complexity of the training phase and get a better reconstruction result, the experimental parameters are set as follows. The topics of pLSA model are chosen 20 and the dictionary size is chosen 500. In order to verify the performance of the proposed method, we compared our proposed method with Bicubic interpolation method, Yang's ScSR method[7], ANR method[15] and JOR method[21].

3.1　Compared our proposed method with other related methods

To illustrate the effectiveness of our method in terms of objective criterion, peak signal to noise ratio (PSNR) and structural similarity index measure (SSIM) are used to evaluate the quality of the reconstruction results. The higher SSIM means that the recovered image is much similar structures to the original image.

In Table 1, the PSNR and SSIM of the recovered images by different methods with magnification factor ×2 and ×3 are displayed. We can see that images reconstructed by our proposed method have higher PSNR and SSIM than other methods. In order to observe the visual effect of the reconstruct HR images, we displayed the reconstructed results of channel 1 image with magnification factor $S = 3$ by different methods in Fig. 3. Obviously, we can see that the HR image recovered by Bicubic interpolation method improves the spatial resolution but produce smooth results. The images recovered HR images by ScSR method and ANR method which can achieve better results than Bicubic interpolation method, but the high frequency detail information is still inadequate. JOR method can get more high frequency detail information but produce reconstruction artifacts obviously. The recovered HR image by our method has adequate detail information and sharper edges. Moreover, our method can reduce the reconstruction artifact.

3.2　Time complexity of reconstruction

In this experiment, we investigate the time complexity of our proposed method and other different super-resolution methods with magnification factor ×2 and ×3. Because the topic dictionaries and the corresponding regression matrices can be pre-computed offline, we only take the reconstruction time into consideration.

We compared our method with ScSR method, ANR method and JOR[21] method. The different reconstruction time is displayed in Table 2. From Table 2, we can see that the ScSR method is really time consuming, ANR method is largely speed up because of avoiding the sparse coding, JOR[21] method also gets a competitive result. However, our proposed method gives a better reconstruction speed-performance trade-off.

3.3　Robustness of the proposed method

Usually, we obtain the images by a variety of sensors. However, there are system internal noise and interference noise whether it is MODIS or other image sensor. Therefore, it is necessary to test the robustness of the super-resolution methods to noise.

In this experiment, we add some Gaussian noise (with zero mean and variance σ) on a test image and demonstrate the performance of our method with different variance of noise. We compared our method with the Bicubic interpolation method, ScSR method, ANR method and JOR[21] method when the noises with $\sigma = 10$, $\sigma = 15$, $\sigma = 20$, $\sigma = 25$ and $\sigma = 30$ are added on the test image.

Table 1　PSNR(dB) and SSIM of the reconstruct images by different methods.

| Factor	Image	Bicubic	ScSR	ANR	JOR	Our method
×2	Channel 1	22.67/0.9665	23.71/0.9739	23.65/0.9727	24.25/0.9767	**24.53/0.9785**
	Channel 2	23.33/0.9696	24.44/0.9767	24.37/0.9775	25.27/0.9813	**25.40/0.9815**
	Channel 3	25.42/0.9369	26.08/0.9465	26.23/0.9732	26.79/0.9787	**26.92/0.9792**
	Channel 4	25.21/0.9433	25.89/0.9522	25.99/0.9740	26.45/0.9746	**26.73/0.9751**
	Channel 7	23.25/0.9680	24.28/0.9750	24.37/0.9775	24.87/0.9795	**25.02/0.9803**
×3	Channel 1	21.06/0.9579	21.78/0.9630	21.79/0.9632	22.01/0.9653	**22.14/0.9656**
	Channel 2	21.64/0.9617	22.42/0.9669	22.42/0.9670	22.96/0.9681	**23.03/0.9688**
	Channel 3	24.22/0.9592	24.73/0.9632	24.72/0.9625	24.93/0.9657	**25.16/0.9662**
	Channel 4	23.97/0.9598	24.50/0.9636	24.50/0.9636	24.86/0.9649	**25.01/0.9654**
	Channel 7	21.62/0.9602	22.37/0.9656	22.39/0.9659	22.60/0.9693	**22.76/0.9680**

OEE | Advances
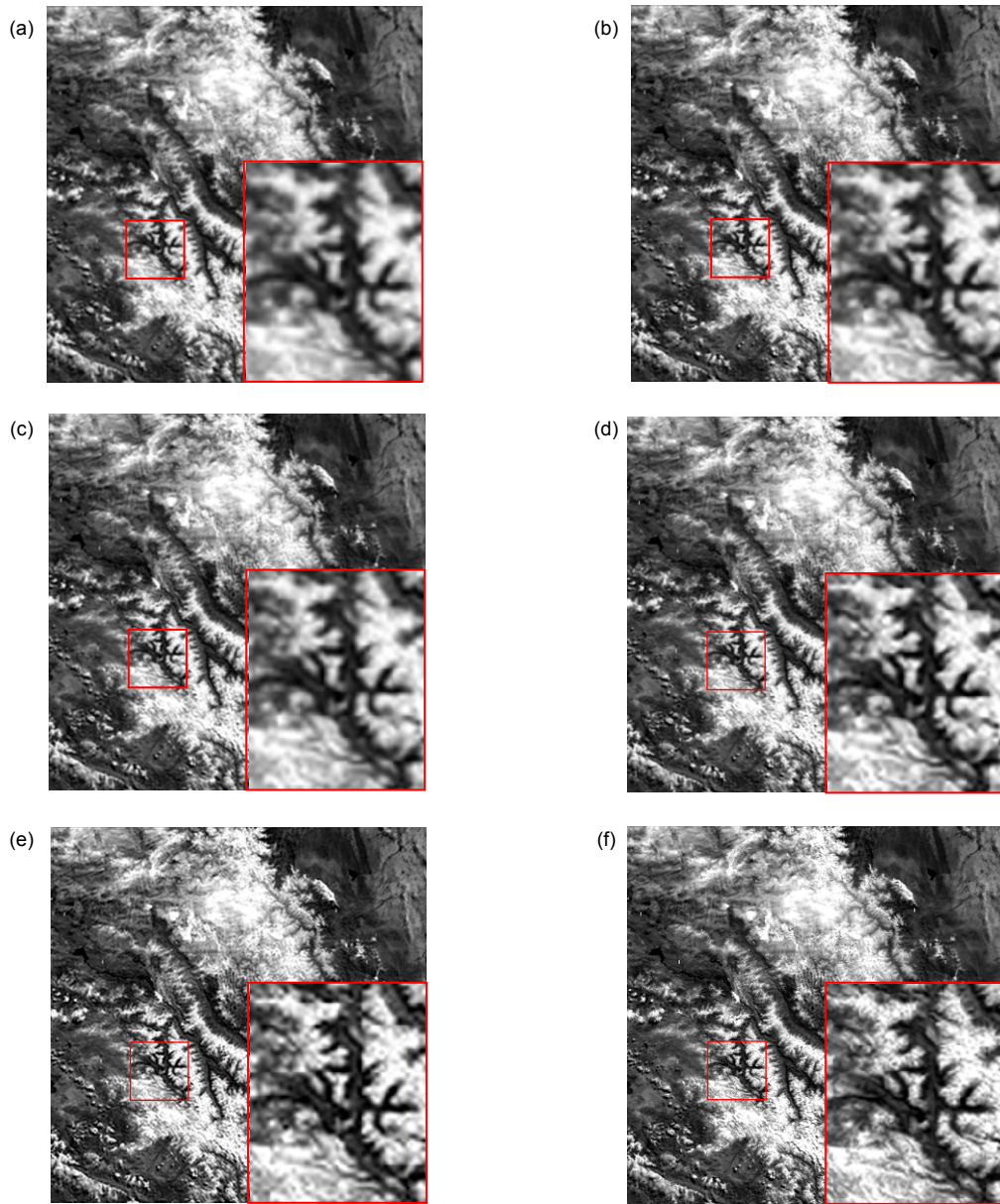
Fig. 3   Visual comparison with different SR results on channel 1 image by different methods (*S*=3). (a) Bicubic method. (b) ScSR method. (c) ANR method. (d) JOR method. (e) Our method. (f) Original image.

Table 2   Consumed time(s) of different methods.

Factor	Method	Channel 1	Channel 2	Channel 3	Channel 4	Channel 7	Average
x2	ScSR	548.87	528.34	536.42	553.60	568.73	547.192
	ANR	3.53	3.26	3.31	3.69	3.51	3.460
	JOR	50.12	47.33	48.45	48.34	52.62	49.372
	Our method	**58.11**	**55.82**	**58.25**	**53.89**	**60.64**	**57.342**
x3	ScSR	537.21	583.82	598.54	612.77	623.28	591.124
	ANR	2.99	2.14	2.89	3.08	3.27	2.874
	JOR	28.34	29.53	30.12	32.21	36.73	31.386
	Our method	**41.93**	**49.01**	**46.25**	**45.25**	**48.54**	**46.196**

The recovered results of different methods are displayed in Fig. 4. From Fig. 4, we can see that when the noise becomes heavier, the PSNR of different methods come heavier, the PSNR of different methods degrades on different levels. However, the PSNR of Bicubic interpolation method, ScSR method, ANR method and JOR[21] method are degraded faster than our method. This means our method has better robustness to noise.

3.4 Discussion of the dictionary size

In this experiment, we will investigate the influence of the dictionary size on the performance of our proposed method. We only vary the dictionary size when investigating the performance of the dictionary size. Fig. 5 displays the PSNR of the reconstruction results when the number of topics was 20 and the dictionary size was from 100 to 1000.

Figure 5(a) displays the results under magnification factor $S = 2$ and Fig. 5(b) displays the results under magnification factor $S = 3$. Obviously, we can see that

the dictionary with more atoms, the reconstruction images get higher PSNR from the curve in Fig. 5. In this experiment, we only displayed the results on channel 1 image because we can get the similar results on other images. Larger dictionary will get better reconstruction results, but the time complexity of dictionary training is also rising fast. In order to make a balance between the time complexity of dictionary training and the reconstruction performance, we implemented our experiments with dictionary size 500 from sections 3.1 to 3.3.

4   Conclusions

In this paper, we proposed an MODIS image super-resolution reconstruction via learned topic dictionaries and regression Matrices. We select the bilateral filter to extract the detail part of image as the object because detail part has a better sparsity. Meanwhile, we focus on the relationships among patches, which results in undesired results. We adopt the pLSA model to discover the
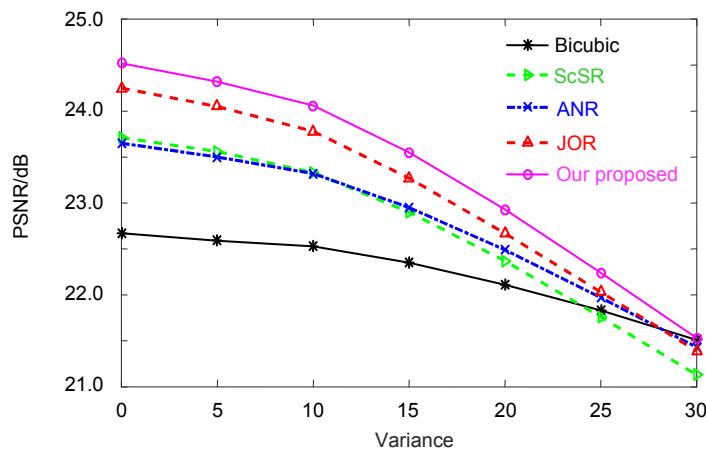

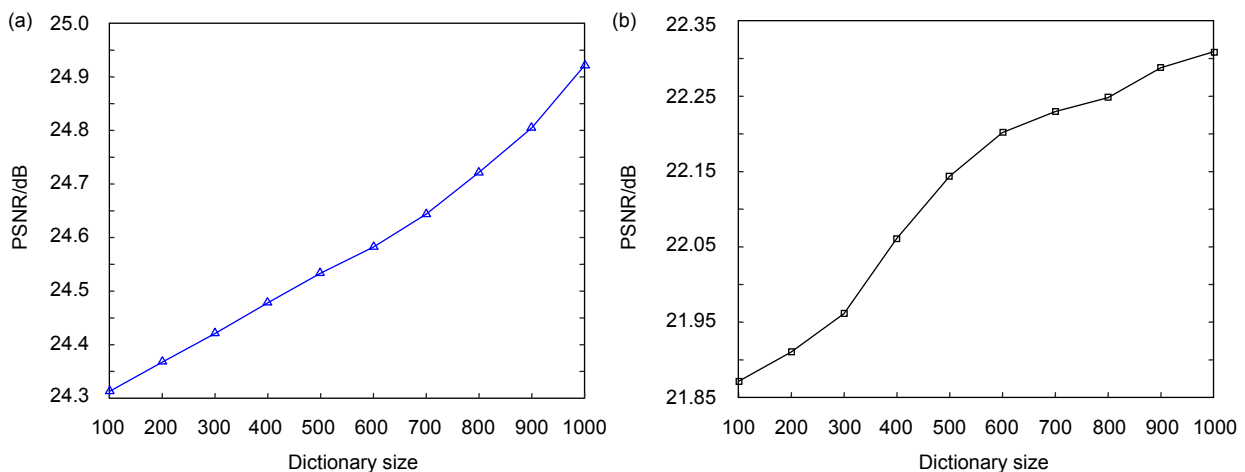Fig. 4   The noise robustness of different methods.

Fig. 5   PSNR (dB) with different dictionary sizes on channel 1 image. (a) S=2. (b) S=3.

semantic relationships about patches. Therefore, we can locally decide which topic it belongs to and the corresponding matrix to predict a possible HR patch for a given LR patch. Our experimental results show that the proposed method, which can greatly reduce the reconstruction artifacts. Moreover, whether from the visual effect or PSNR, SSIM and the robustness to noise, our method performs superior to other three methods.

Acknowledgements

References

1   Farsiu S, Robinson M D, Elad M, *et al*. Fast and robust multi-frame super resolution[J]. *IEEE Transactions on Image Processing*, 2004, **13**(10): 1327–1344.

2   Lu Yao, Inamura M. Spatial resolution improvement of remote sensing images by fusion of subpixel-shifted multi-observation images[J]. *International Journal of Remote Sensing*, 2003, **24**(23): 4647–4660.

3   Aumann H H, Chahine M T, Gautier C, *et al*. AIRS/AMSU/HSB on the Aqua mission: design, science objectives, data products, and processing systems[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2003, **41**(2): 253–264.

4   Chan T F, Ng M K, Yau A C, *et al*. Superresolution image reconstruction using fast inpainting algorithms[J]. *Applied and Computational Harmonic Analysis*, 2007, **23**(1): 3–24.

5   Keys R. Cubic convolution interpolation for digital image processing[J]. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 2003, **29**(6): 1153–1160.

6   Yang Shuyuan, Wang Min, Chen Yiguang, *et al*. Single-image super-resolution reconstruction via learned geometric dictionaries and clustered sparse coding[J]. *IEEE Transactions on Image Processing*, 2012, **21**(9): 4016–4028.

7   Yang Jianchao, Wright J, Huang T, *et al*. Image super-resolution as sparse representation of raw image patches[C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage AK, USA, 2008: 1–8.

8   Yang Jianchao, Wright J, Huang T S, *et al*. Image super-resolution via sparse representation[J]. *IEEE Transactions on Image Processing*, 2010, **19**(11): 2861–2873.

9   Zhou Ying, Fu Randi, Yan Wen, *et al*. A method of infrared nephogram super-resolution based on structural group sparse representation[J]. *Opto-Electronic Engineering,* 2016, **43**(12): 126–132.

10   Chang Hong, Yeung D Y, Xiong Yimin M. Super-resolution through neighbor embedding[C]//*Proceedings of the 2004 IEEE Computer Society Computer Vision and Pattern Recognition*, Washington DC, USA, 2004: 275–282.

11   Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding[J]. *Science*, 2000, **290**(5500): 2323–2326.

12   Donoho D L. Compressed sensing[J]. *IEEE Transactions on Information Theory*, 2012, **52**(4): 1289–1306.

13   Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[C]//*Proceedings of 7th International Conference on Curves and Surfaces*, Avignon, France, 2010: 711–730.

14   Aharon M, Elad M, Bruckstein A. *rmK*-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. *IEEE Transactions on Signal Processing*, 2006, **54**(11): 4311–4322.

15   Timofte R, De V, Gool L V. Anchored neighborhood regression for fast example-based super-resolution[C]//*Proceedings of IEEE International Conference on Computer Vision*, Sydney NSW, Australia, 2013: 1920–1927.

16   Hofmann T. Probabilistic latent semantic analysis[C]//*Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*, Stockholm, Sweden, 1999: 289–296.

17   Zhao Yao, Chen Qian, Sui Xiubao, *et al*. A novel infrared image super-resolution method based on sparse representation[J]. *Infrared Physics & Technology*, 2015, **71**: 506–513.

18   Purkait P, Chanda B. Image upscaling using multiple dictionaries of natural image patches[C]//*Proceedings of Asian Conference on Computer Vision*, Daejeon, Korea, 2012: 284–295.

19   Rubinstein R, Zibulevsky M, Elad M. Efficient Implementation of the K-SVD Algorithm Using Batch Orthogonal Matching Pursuit[R]. Technion-Computer Science Department, Technical Report CS, 2008.

20   Timofte R, van Gool L. Adaptive and Weighted Collaborative Representations for image classification[J]. *Pattern Recognition Letters*, 2014, **43**: 127–135.

21   Dai D, Timofte R, Van Gool L. Jointly optimized regressors for image super-resolution[J]. *Computer Graphics Forum*, 2015, **34**(2): 95–104.